Convex Optimization
Lecture Notes for EE 227BT
Draft, Fall 2013

Laurent El Ghaoui

August 29, 2013

Contents

Chapter 1

Introduction

1.1 Optimization problems

1.1.1 The model

The mathematical programming model is of the form

$$p^* := \min_x \ f_0(x) \text{ subject to } f_i(x) \leq 0, \ \ i = 1, \ldots, m. \tag{1.1}$$

- $x \in \mathbf{R}^n$ is the *decision variable*
- $f_0 : \mathbf{R}^n \to \mathbf{R}$ is the *objective* function
- $f_i : \mathbf{R}^n \to \mathbf{R}$, $i = 1, \ldots, m$ represent the *constraints*
- p^* is the *optimal value*

The term "programming" (or "program") does not refer to a computer code. It is used mainly for historical purposes. A more rigorous (but less popular) term is "optimization problem". The term "subject to" is often replaced by a colon.

The set

$$\mathcal{D} = \{x \in \mathbf{R}^n \ : \ f_i(x) \leq 0, \ \ i = 1, \ldots, m\}$$

is called the *feasible set*. Any $x \in \mathcal{D}$ is called feasible (with respect to the specific optimization problem at hand).

Sometimes, the model is described in terms of the feasible set, as follows:

$$\min_{x \in \mathcal{D}} \ f_0(x).$$

Also, sometimes a maximization problem is considered:

$$\max_{x \in \mathcal{D}} \ f_0(x).$$

Finally, it may be useful to distinguish between "structure" constraints (such as non-negativity constraints on variables) and constraints involving problem data, with the notation

$$\min_{x \in \mathcal{X}} \ f_0(x) \text{ subject to } f_i(x) \leq 0, \ \ i = 1, \ldots, m$$

where $\mathcal{X} \subseteq \mathbf{R}^n$ describes the "structure" constraints.

1.1.2 Examples

- Least-squares problem:

$$\min_x \ \|Ax - b\|_2^2$$

  where $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$, and $\|\cdot\|_2$ denotes the Euclidean norm.

  This problem arises in many situations, for example in statistical estimation problems such as linear regression. The problem dates back many years, at least to Gauss (1777-1855), who solved it to predict the trajectory of the planetoid Ceres.

- Linear programming problem:

$$\min \ c^T x \ : \ a_i^T x \leq b_i, \ \ i = 1, \ldots, m,$$

  where $c \in \mathbf{R}^n$, $a_i \in \mathbf{R}^n$, $b_i \in \mathbf{R}$ $(i = 1, \ldots, m)$. This corresponds to the case where the functions f_i $(i = 0, \ldots, m)$ in (1.1) are all affine (that is, linear plus a constant term).

  This problem was introduced by Dantzig in the 40's in the context of logistical problems arising in military operations. This model of computation is perhaps the most widely used optimization problem today.

- Quadratic programming problem:

$$\min \ \|x\|_2^2 + c^T x \ : \ a_i^T x \leq b_i, \ \ i = 1, \ldots, m,$$

  which can be thought of as a generalization of both the least-squares and linear programming problems.

  QP's are popular in many areas, such as finance, where the linear term in the objective refers to the expected negative return on an investment, and the squared term corresponds to the risk (or variance of the return).

  This model was introduced by Markowitz (who was a student of Dantzig) in the 50's, to model investment problems. Markowitz won the Nobel prize in Economics in 1990, mainly for this work.

1.1.3 Solution

The *optimal set* of problem (1.1) is defined as the set of *feasible* points x^* such that $p^* = f_0(x^*)$:

$$\mathcal{X}_{\rm opt} := \{x \in \mathbf{R}^n \ : \ f_i(x) \leq 0, \ \ i = 1, \ldots, m, \ \ p^* = f_0(x)\}\,.$$

The ϵ-suboptimal set is defined as

$$\mathcal{X}_\epsilon := \{x \in \mathbf{R}^n \ : \ f_i(x) \leq 0, \ \ i = 1, \ldots, m, \ \ f_0(x) \leq p^* + \epsilon\}\,.$$

(With our notation, $\mathcal{X}_0 = \mathcal{X}_{\rm opt}$.)

A point z is *locally optimal* if there is a value $R > 0$ such that z is optimal for problem

$$\min_x \ f_0(x) \text{ subject to } f_i(x) \leq 0, \ \ i = 1, \ldots, m, \ \ \|z - x\|_2 \leq R.$$

In other words, x minimizes f_0, but only for nearby points on the feasible set.

1.2 Convex optimization problems

1.2.1 Convexity

A set $\mathcal{C}$ is convex if it contains the line segments between any two of its points:

$$\forall\, x, y \in \mathcal{C},\ \ \forall\, \lambda \in [0,1],\ \ \lambda x + (1-\lambda) y \in \mathcal{C}.$$

A function $f : \mathbf{R}^n \to \mathbf{R}$ is *convex* if

$$\forall\, x, y \in \mathbf{R}^n,\ \ \forall\, \lambda \in [0,1],\ \ f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y).$$

In other words, the graph of the function is always below the chord joining any two points. That is, a function is convex if and only if its *epigraph*

$$\mathbf{epi} f := \left\{ (x,t) \in \mathbf{R}^{n+1} \ :\ t \geq f(x) \right\}$$

is convex. (Check this statement.)

The optimization problem (1.1) is convex if every function involved $f_0, f_1, \ldots, f_m$, is convex.

The examples presented in section (1.1.2) are all convex. Examples of non-convex problems include *combinatorial optimization* problems, where (some if not all) variables are constrained to be boolean, or integers. (Such problems arise for example when discrete choices are to be made, such as in crew asssignment in the airline industry.)

1.2.2 Complexity

In this course, complexity of an optimization problem refers to the difficulty of solving the problem on a computer. At this stage we do not define this notion precisely.

The complexity of an optimization problem depends on its structure. Two seemingly similar problem may require a widely different computational effort to solve. Some problems are "NP-hard", which roughly means that they cannot be solved in reasonable time on a computer.

As an example, the quadratic programming problem seen above is "easy" to solve, however the apparently similar problem

$$\min\ c^T x - \|x\|_2^2\ :\ a_i^T x \leq b_i,\ \ i = 1, \ldots, m,$$

is NP-hard.

In the early days of optimization, it was thought that linearity was what distinguished a hard problem from an easy one. Today, it appears that convexity is the relevant notion. Roughly speaking, a convex problem is easy. In this course, we will refine this statement.

1.2.3 A brief history of convex optimization

Theory:

- 19-th century: optimization models are used mostly in physics, with the concept of energy as the objective function. No attempt (with the notable exception of Gauss' algorithm for least-squares) is made to actually solve these problems numerically.
- 1900-70: a great effort in made in mathematics to build the theory of optimization. The emphasis is on convex analysis, which allows to describe the optimality conditions of a convex problem.

Algorithms:

- 1947: simplex algorithm for linear programming (Dantzig).
- 1960s: early interior-point methods (Fiacco & McCormick, Dikin, ...).
- 1970s: ellipsoid method and other subgradient methods
- 1980s: polynomial-time interior-point methods for linear programming (Karmarkar 1984).
- late 1980s–now: polynomial-time interior-point methods for nonlinear convex optimization (Nesterov & Nemirovski 1994).

Applications (of convex optimization):

- before 1990: mostly in operations research; few in engineering (except least-squares); statistics is a big user of nonlinear optimization methods such as Newton-Raphson.
- since 1990: many new applications in engineering (control, signal processing, communications, circuit design, ...); new problem classes (semidefinite and second-order cone programming, robust optimization)

1.3 Course objectives

The course will emphasize models, not algorithms. It will cover a specific class of optimization models, based on the notion of convexity. We will briefly cover some algorithms for solving convex problems. We will also study robust optimization, which is a method to handle uncertainty in decision-making.

Chapter 2

Linear Algebra Review

2.1 Vectors

2.1.1 Basics

Independence. A set of vectors $x_i \in \mathbf{R}^n$, $i = 1, \ldots, m$ is said to be *independent* if and only if the following condition on a vector $\lambda \in \mathbf{R}^m$:

$$\sum_{i=1}^{m} \lambda_i x_i = 0$$

implies $\lambda = 0$. This means that no vector in the set can be expressed as a linear combination of the others.

Subspace, span. A subspace of $\mathbf{R}^n$ is a subset that is closed under addition and scalar multiplication. As an example, the *span* of a set of vectors $x_i \in \mathbf{R}^n$, $i = 1, \ldots, m$ is defined as

$$\text{span}(x_1, \ldots, x_m) := \left\{ \sum_{i=1}^{m} \lambda_i x_i \ : \ \lambda \in \mathbf{R}^m \right\}.$$

Basis. A basis of $\mathbf{R}^n$ is a set of n independent vectors. The basis of a given subspace $\mathcal{L} \subseteq \mathbf{R}^n$ is any independent set of vectors whose span is $\mathcal{L}$. The number of vectors in the basis is actually independent of the choice of the basis (for example, in $\mathbf{R}^3$ you need two independent vectors to describe a plane containing the origin). This number is called the *dimension* of $\mathcal{L}$.

2.1.2 Scalar product and norms

Scalar product. The scalar product (or, dot product) between two vectors $x, y \in \mathbf{R}^n$ is defined as the scalar $x^T y = \sum_{i=1}^{n} x_i y_i$. More generally, an inner product on $\mathbf{R}^n$ is a bilinear function $\langle \cdot, \cdot \rangle$ that satisfies the properties of symmetry (with respect to a swap in the two arguments), and positive-definiteness (that is, $\langle x, x \rangle$ is always non-negative, and zero only when $x = 0$). An example of an inner product is the weighted dot product

$$\langle x, y \rangle_\sigma := \sum_{i=1}^{m} \sigma_i^2 x_i y_i, \tag{2.1}$$

where $\sigma \in \mathbf{R}^n$, $\sigma \neq 0$ is given.

Vector norms. A function $f : \mathbf{R}^n \to \mathbf{R}$ is a norm on $\mathbf{R}^n$ if the following three conditions are met:

1. f is convex.

2. f is positively homogeneous, meaning that $f(\alpha x) = \alpha f(x)$ for every $x \in \mathbf{R}^n$ and $\alpha \in \mathbf{R}_+$.

3. f is positive-definite: for every $x \in \mathbf{R}^n$, $f(x) = 0$ implies $x = 0$.

Together, the first two conditions are equivalent to the *triangle inequality*:

$$\forall\, x, y \in \mathbf{R}^n \;:\; f(x+y) \leq f(x) + f(y).$$

Often, norms are denoted $\|\cdot\|$.

Popular norms. There are three very popular norms for a vector $x \in \mathbf{R}^n$:

- The Euclidean norm is $\|x\|_2 := \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{x^T x}$, which corresponds to the usual notion of distance in two or three dimensions.

- The l_1-norm, or Manhattan distance, is $\|x\|_1 = \sum_{i=1}^n |x_i|$. The norm corresponds to the distance travelled on a rectangular grid (such as Manhattan) to go from one point to another.

- The l_∞-norm is given by $\|x\|_\infty := \max_{1 \leq i \leq n} |x_i|$.

The l_p-norm is a class that includes the previous ones (in an asymptotic sense in the case of the l_∞ norm), and is defined as

$$\|x\|_p := \left(\sum_{i=1}^p |x_i|^p \right)^{1/p},$$

where $p \geq 1$.

There are many other norms that are important or interesting in some applications. For example, for $k \in \{1, \ldots, n\}$ we can define

$$\|x\|_{1,k} := \sum_{i=1}^k |x|_{[i]}$$

where for every i, $|x|_{[i]}$ is the i-th largest absolute value of elements of x. The norm is a kind of mixture between the l_1- and l_∞-norms, respectively obtained upon setting $k = n$ and $k = 1$.

Finally, any scalar product $\langle \cdot, \cdot \rangle$ generates a norm, defined as $\|x\| := \sqrt{\langle x, x \rangle}$. For example, the Euclidean norm is generated by the ordinary scalar product. Another example is the norm induced by the inner product defined in (2.1), which is the weighted Euclidean norm

$$\|x\| = \sqrt{\sum_{i=1}^n \sigma_i^2 x_i^2}.$$

Cauchy-Schwartz inequalities, angles, dual norm. The Cauchy-Schwartz inequality states that

$$\forall\, x, y \in \mathbf{R}^n \;:\; x^T y \leq \|x\|_2 \cdot \|y\|_2.$$

When none of the vectors involved is zero, we can define the corresponding angle as θ such that

$$\cos\theta = \frac{x^T y}{\|x\|_2 \|y\|_2}.$$

(The notion generalizes the usual notion of angle between two directions in two dimensions.)

Cauchy-Schwartz inequalities can be obtained for norms other than the Euclidean. For example,

$$\forall\, x, y \in \mathbf{R}^n \;:\; x^T y \leq \|x\|_\infty \cdot \|y\|_1.$$

More generally, to any norm $\|\cdot\|$ we can associate a *dual norm*, usually denoted $\|\cdot\|_*$, and defined as

$$\|y\|_* := \max_x \; x^T y \;:\; \|x\| \leq 1.$$

(Check this is indeed a norm.) By construction, the norm $\|\cdot\|$ and its dual satisfy the (generalized) Cauchy-Schwartz inequality

$$\forall\, x, y \in \mathbf{R}^n \;:\; x^T y \leq \|x\| \cdot \|y\|_*.$$

In this setting, the Euclidean norm is its own dual; and the l_1- and l_∞-norms are dual of each other.

Orthogonal basis. A basis $(u_i)_{i=1}^n$ is said to be *orthogonal* if $u_i^T u_j = 0$ if $i \neq j$. If in addition, $\|u_i\|_2 = 1$, we say that the basis is *orthonormal*.

2.2 Matrices

2.2.1 Basics

Matrices (in say, $\mathbf{R}^{m\times n}$) can be viewed simply as a collection of vectors of same size. Alternatively, a matrix can be see as a (linear) operator from the "input" space $\mathbf{R}^n$ to the "output" space $\mathbf{R}^m$. Both points of view are useful.

Transpose, trace and scalar product. The transpose of a matrix A is denoted by A^T, and is the matrix with (i, j) element A_{ji}, $i = 1, \ldots, m$, $j = 1, \ldots, n$.

The *trace* of a square $n \times n$ matrix A, denoted by $\mathbf{Tr}\, A$, is the sum of its diagonal elements: $\mathbf{Tr}\, A = \sum_{i=1}^n A_{ii}$.

We can define the scalar product between two $m \times n$ matrices A, B via

$$\langle A, B\rangle = \mathbf{Tr}\, A^T B = \sum_{i=1}^m \sum_{j=1}^m A_{ij} B_{ij}.$$

In this definition, both A, B are viewed as long vectors with all the columns stacked on top of each other, and the scalar product is the ordinary scalar product between the two vectors.

Range, nullspace, rank. The range of a $m \times n$ matrix A is defined as the following subset of $\mathbf{R}^m$:

$$\mathcal{R}(A) := \{Ax \ : \ x \in \mathbf{R}^n\}.$$

The nullspace of A is given by

$$\mathcal{N}(A) := \{x \in \mathbf{R}^n \ : \ Ax = 0\}.$$

The rank of a matrix A is the dimension of its range; it is also the rank of A^T. Alternatively, it is equal to n minus the dimension of its nullspace. A basic result of linear algebra states that any vector in $\mathbf{R}^n$ can be decomposed as $x = y + z$, with $y \in \mathcal{N}(A)$, $z \in \mathcal{R}(A^T)$, and z, y are orthogonal. (One way to prove this is via the singular value decomposition, seen later.)

The notions of range, nullspace and rank are all based on viewing the matrix as an operator.

Orthogonal matrices. A square, $n \times n$ matrix $U = [u_1, \ldots, u_n]$ is orthogonal if its columns form an orthonormal basis (note the unfortunate wording). The condition $u_i^T u_j = 0$ if $i \neq j$, and 1 otherwise, translates in matrix terms as $U^T U = I_n$ with I_n the $n \times n$ identity matrix.

Orthogonal matrices are sometimes called rotations. Indeed, they do not change the Euclidean norm of the input: for every $x \in \mathbf{R}^n$, we have $\|Ux\|_2 = \|x\|_2$ (why?).

2.2.2 Matrix norms

There are many ways to define the norm of a matrix $A \in \mathbf{R}^{m \times n}$.

A first class of matrix norms, which can be called *vector-based,* can be derived by simply collecting the elements of the matrix into a big vector, and defining the matrix norm to be the norm of that vector. A popular choice in this class is the *Frobenius* norm, which corresponds to the Euclidean norm of the vector formed with its elements:

$$\|A\|_F := \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij}^2}.$$

Another class of matrix norm can be obtained as *induced* by a vector norm. Specifically, let $\|\cdot\|_{\text{in}}$, $\|\cdot\|_{\text{out}}$ be two vector norms defined on $\mathbf{R}^n$ and $\mathbf{R}^m$, respectively. Then we define the norm of a $m \times n$ matrix A as

$$\|A\| := \max_x \ \|Ax\|_{\text{out}} \ : \ \|x\|_{\text{in}} \leq 1.$$

It turns out that the above indeed defines a matrix norm. This class of norms views A not as a vector, but as a linear operator, and the norm measures the maximum norm (measured with the output norm $\|\cdot\|_{\text{out}}$) that the operator can achieve with bounded inputs (with bounds measured via the "input" norm $\|\cdot\|_{\text{in}}$).

One popular choice corresponds to the case when both input and output norms are Euclidean. This norm is called the *largest singular value* norm, for reasons visited later.

Some norms are both vector-based and induced. The Frobenius norm is not induced; and the largest singular value norm is not vector-based.

2.2.3 Matrix description of subspaces

Linear and affine subspace. A subspace in $\mathbf{R}^n$ can always be described as the nullspace of a matrix A:

$$\mathcal{L} = \{x \in \mathbf{R}^n \ : \ Ax = 0\}.$$

The dimension of $\mathcal{L}$ is the rank of the matrix A. The subspace above is simply the span of the columns of A.

A subset of the form

$$\mathcal{L} = \{x \in \mathbf{R}^n \ : \ Ax = b\},$$

with $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$, is referred to as an *affine subspace.*

Hyperplanes. A hyperplane in $\mathbf{R}^n$ is a set described by one affine constraint. Hence, it is an affine subspace of dimension $n-1$. It can be described by one vector $a \in \mathbf{R}^n$ and one scalar b:

$$\mathcal{H} = \left\{x \in \mathbf{R}^n \ : \ a^T x = b\right\}.$$

2.2.4 Singular value decomposition

The singular value decomposition states that any matrix $A \in \mathbf{R}^{m\times n}$ can be expressed as

$$A = U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^T,$$

with $U \in \mathbf{R}^{m\times m}$, $V \in \mathbf{R}^{n\times n}$, U, V orthogonal, and $\Sigma = \mathbf{diag}(\sigma_1, \ldots, \sigma_r)$ contain the *singular values* of A. The number $r \leq \min(m, n)$ is the rank of A, and equals the dimension of its range.

The largest singular value of A can be characterized as

$$\sigma_{\max}(A) = \max_{1\leq i\leq r} \sigma_i = \max_x \|Ax\|_2 \ : \ \|x\|_2 = 1.$$

(Try to prove this.) As mentioned before, the largest singular value is a matrix norm.

Due to the orthogonality of the matrices U, V, the SVD is especially useful in connection with the Euclidean norm, or to analyze linear equations (we can extract information about rank, nullspace and range from the SVD).

2.3 Symmetric Matrices

2.3.1 Definition and examples

Definition. A square matrix $A \in \mathbf{R}^{n\times n}$ is *symmetric* if and only if $A = A^T$. The set of symmetric $n \times n$ matrices is denoted $\mathcal{S}^n$.

Examples. Perhaps the simplest example of symmetric matrices is the class of diagonal matrices, which are non-zero only on their diagonal. If $\lambda \in \mathbf{R}^n$, we denote by $\mathbf{diag}(\lambda_1, \ldots, \lambda_n)$, or $\mathbf{diag}(\lambda)$ for short, the $n \times n$ (symmetric) diagonal matrix with λ on its diagonal.

Another important case of a symmetric matrix is of the form uu^T, where $u \in \mathbf{R}^n$. The matrix has elements $u_i u_j$, and is symmetric. Such matrices are called *dyads.* If $\|u\|_2 = 1$, then the dyad is said to be normalized.

A symmetric matrix is a way to describe a weighted, undirected graph: each edge in the graph is assigned a weight A_{ij}. Since the graph is undirected, the edge weight is independent of the direction (from i to j or vice-versa). Hence, A is symmetric.

Another interesting special case of a symmetric matrix is the Jacobian of a function at a given point, which is the matrix containing the second derivatives of the function. Here, we invoke the fact that the second-derivative is independent of the order in which derivatives are taken.

Finally, any quadratic function $q : \mathbf{R}^n \to \mathbf{R}$ can be written as

$$q(x) = \begin{pmatrix} x \\ 1 \end{pmatrix}^T A \begin{pmatrix} x \\ 1 \end{pmatrix},$$

for an appropriate symmetric matrix $A \in \mathcal{S}^{(n+1)}$. If q is a quadratic form (meaning that there are no linear or constant terms in it), then we can write $q(x) = x^T Ax$ where now $A \in \mathcal{S}^n$.

2.3.2 Eigenvalue decomposition

A fundamental result of linear algebra states that any symmetric matrix can be decomposed as a weighted sum of normalized dyads that are orthogonal to each other.

Precisely, for every $A \in \mathcal{S}^n$, there exist numbers $\lambda_1, \ldots, \lambda_n$ and an orthonormal basis $(u_1, \ldots, u_n)$, such that

$$A = \sum_{i=1}^{n} \lambda_i u_i u_i^T.$$

In a more compact matrix notation, we have $A = U\Lambda U^T$, with $\Lambda = \mathbf{diag}(\lambda_1, \ldots, \lambda_n)$, and $U = [u_1, \ldots, u_n]$.

The numbers $\lambda_1, \ldots, \lambda_n$ are called the eigenvalues of A, and are the roots of the characteristic equation

$$\det(\lambda I - A) = 0,$$

where I_n is the $n \times n$ identity matrix. For arbitrary square matrices, eigenvalues can be complex. In the symmetric case, the eigenvalues are always real. Up to a permutation, eigenvalues are unique, in the sense that there are only n (possibly distinct) solutions to the above equation.

The vectors u_i, $i = 1, \ldots, n$, are called the (normalized) *eigenvectors* of A. In contrast with eigenvalues, there is no unicity property here. For example, the identity matrix has any (unit-norm) vector as eigenvector. However, if all the eigenvalues are distinct, then eigenvectors are unique (up to a change in sign).

It is interesting to see what the eigenvalue decomposition of a given symmetric matrix A tells us about the corresponding quadratic form, $q_A(x) := x^T Ax$. With $A = U\Lambda U^T$, we have

$$q_A(x) = (U^T x)^T \Lambda (U^T x) = \sum_{i=1}^{n} \lambda_i (u_i^T x)^2.$$

The eigenvalue decomposition thus corresponds to the decomposition of the corresponding quadratic form into a sum of squares.

2.3.3 Positive semi-definite matrices

Definition. A matrix $A \in \mathcal{S}^n$ is said to be *positive-definite* (resp. *positive semi-definite*) if and only if all the eigenvalues are positive (resp. non-negative). We use the acronyms PD and PSD for these properties. The set of $n \times n$ PSD matrices is denoted $\mathcal{S}^n_+$, while that of PD matrices is written $\mathcal{S}^n_{++}$. Often, we use the notation $A \succeq 0$ (resp. $A \succ$) for the PSD (resp. PD) property.

In terms of the associated quadratic form $q_A(x) = x^T Ax$, the interpretation is as follows. A matrix A is PD if and only if q_A is a positive-definite function, that is, $q_A(x) = 0$ if and only if $x = 0$. Indeed, when $\lambda_i > 0$ for every i, then the condition

$$q_A(x) = \sum_{i=1}^{n} \lambda_i (u_i^T x)^2 = 0$$

trivially implies $u_i^T x = 0$ for every i, which can be written as $Ux = 0$. Since U is orthogonal, it is invertible, and we conclude that $x = 0$. Thus, to any PD matrix A, we can associate a norm, $\|x\|_A := \sqrt{x^T Ax}$.

Square root and Cholesky decomposition. For PD matrices, we can generalize the notion of ordinary square root of a non-negative number. Indeed, if A is PSD, there exist a unique PD matrix, denoted $A^{1/2}$, such that $A = (A^{1/2})^2$. If A is PD, then so is its square root.

Any PSD matrix can be written as a product $A = LL^T$ for an appropriate matrix L. The decomposition is not unique, and $R = A^{1/2}$ is only a possible choice. If A is positive-definite, then we can choose L to be lower triangular, and invertible. The decomposition is then known as the Cholesky decomposition. The corresponding weighted norm $\|x\|_A$ mentioned above is then simply the Euclidean norm of $L^T x$.

Examples and interpretations. A well-known example of a PSD matrix is the covariance matrix associated with a random variable in $\mathbf{R}^n$. This matrix is defined as

$$\Sigma = \mathbf{E}(x - \hat{x})(x - \hat{x})^T,$$

where $\hat{x} := \mathbf{E}\, x$, and $\mathbf{E}$ denotes the expectation operator associated with the distribution of the random variable x.

Another important example is geometric. For a PD matrix P, and vector $\hat{x}$, the set

$$\mathcal{E}(\hat{x}, P) := \left\{ x \ : \ (x - \hat{x})^T P^{-1} (x - \hat{x}) \leq 1 \right\}$$

is an ellipsoid, with center $\hat{x}$. Its principal axes are given by the orthogonal basis that diagonalizes P, and the semi-axis lengths are the eigenvalues. (Check what happens when P is proportional to the identity.) If P is factored as $P = LL^T$ for some (lower-triangular) matrix L, then the ellipsoid can be interpreted as the affine transformation of a unit Euclidean ball:

$$\mathcal{E}(\hat{x}, P) = \left\{ \hat{x} + Lu \ : \ \|u\|_2 \leq 1 \right\}.$$

Chapter 3

Convex Optimization Problems

3.1 Convex Sets

3.1.1 Definition

A subset $\mathcal{C}$ of $\mathbf{R}^n$ is convex if and only if it contains the line segment between any two points in it:

$$\forall\, x_1, x_2 \in \mathcal{C},\ \ \forall\, \theta_1 \geq 0,\ \theta_2 \geq 0,\ \theta_1 + \theta_2 = 1\ :\ \theta_1 x_1 + \theta_2 x_2 \in \mathcal{C}.$$

Some important special cases of convex sets are the following.

- The set is said to be an *affine subspace* if it contains the entire line passing through any two points. This corresponds to the condition above, with θ_1, θ_2 arbitrary. Subspaces and affine subspaces are convex.
- The set is said to be a *convex cone* if the condition above holds, but with the restriction $\theta_1 + \theta_2 = 1$ removed.

3.1.2 Examples

- The *convex hull* of a set of points $\{x_1, \ldots, x_m\}$ is defined as

  $$\mathbf{Co}(x_1, \ldots, x_m) := \left\{ \sum_{i=1}^m \lambda_i x_i\ :\ \lambda \in \mathbf{R}^m_+,\ \ \sum_{i=1}^m \lambda_i = 1 \right\},$$

  and is convex. The *conic hull:*

  $$\left\{ \sum_{i=1}^m \lambda_i x_i\ :\ \lambda \in \mathbf{R}^m_+ \right\}$$

  is a convex cone.
- For $a \in \mathbf{R}^n$, and $b \in \mathbf{R}$, the hyperplane $\mathcal{H} = \{x\ :\ a^T x = b\}$ is affine. The *half-space* $\{x\ :\ a^T x \leq b\}$ is convex.
- For a square, non-singular matrix $R \in \mathbf{R}^{n \times n}$, and $x_c \in \mathbf{R}^n$, the *ellipsoid* $\{x_c + Ru\ :\ \|u\|_2 \leq 1\}$ is convex. (The center of the epllipsoid is x_c, and you can think of R as the "radius".) With $P = RR^T$, we can describe the ellipsoid as

  $$\left\{x\ :\ (x - x_c)^T P^{-1} (x - x_c) \leq 1\right\}.$$

- A *polyhedron* is a set described by a finite number of affine inequalities and equalities:
$$\mathcal{P} = \{x \ : \ Ax \leq b, \ \ Cx = d\}\,,$$
where A, C are matrices, b, d are vectors, and inequalities are understood component-wise. Sometimes bounded polyhedra are referred to as polytopes.

- The *probability simplex*
$$\left\{p \in \mathbf{R}^n_+ \ : \ \sum_{i=1}^n p_i = 1\right\}$$
is a special case of a polyhedron, and is useful to describe discrete probabilities.

- The *second-order cone*
$$\left\{(x,t) \in \mathbf{R}^{n+1} \ : \ t \geq \|x\|_2\right\} \tag{3.1}$$
is a convex cone. It is sometimes called "ice-cream cone", for obvious reasons. (We will prove the convexity of this set later.)

- The positive semi-definite cone
$$\mathcal{S}^n_+ := \left\{X = X^T \in \mathbf{R}^{n\times n} \ : \ X \succeq 0\right\}$$
is a convex cone. (Again, we will prove the convexity of this set later.)

3.1.3 Support and indicator functions

For a given set S, the function
$$\phi_S(x) := \max_{u\in S} \ x^T u$$
is called the *support function* of S. If S is the unit ball for some norm: $S = \{u : \|u\| \leq 1\}$, then the support function of S is the dual norm. Another important function associated with S is the indicator function
$$I_S(x) = \begin{cases} 0 & x \in S, \\ +\infty & x \notin S. \end{cases}$$

3.1.4 Operations that preserve convexity

Two important operations that preserve convexity are:

- *Intersection:* the intersection of a (possibly infinite) family of convex sets is convex. We can use this property to prove that the semi-definite cone $\mathcal{S}^n_+$ is convex, since
$$\mathcal{S}^n_+ = \left\{X = X^T \in \mathbf{R}^{n\times n} \ : \ \forall\, z \in \mathbf{R}^n, \ \ z^T X z \geq 0\right\},$$
from which we see that the set is the intersection of the subspace of symmetric matrices with a set described by an infinite number of linear inequalities of the form $z^T X z \geq 0$, indexed by $z \in \mathbf{R}^n$. Likewise, the second-order cone defined in (3.1) is convex, since the condition $t \geq \|x\|_2$ is equivalent to the infinite number of affine inequalities $t \geq u^T x$, $\|u\|_2 \leq 1$.

- *Affine transformation:* If a function is affine (that is, it is the sum of a linear function and a constant), and $\mathcal{C}$ is convex, then the set
$$f(\mathcal{C}) := \{f(x) \ : \ x \in \mathcal{C}\}$$
is convex. A particular example is projection on a subspace, which preserves convexity.

3.1.5 Separation theorems

There are many versions of separation theorems in convex analysis. One of them is the separating hyperplane theorem:

Theorem 1 (Separating hyperplane) *If $\mathcal{C}$, $\mathcal{D}$ are two convex subsets of $\mathbf{R}^n$ that do not intersect, then there is an hyperplane that separates them, that is, there exit $a \in \mathbf{R}^n$, $a \neq 0$, and $b \in \mathbf{R}$, such that $a^T x \leq b$ for every $x \in \mathcal{C}$, and $a^T x \geq b$ for every $x \in \mathcal{D}$.*

Another result involves the separation of a set from a point on its boundary:

Theorem 2 (Supporting hyperplane) *If $\mathcal{C} \subseteq \mathbf{R}^n$ is convex and non-empty, then for any x_0 at the boundary of $\mathcal{C}$, there exist a supporting hyperplane to $\mathcal{C}$ at x_0, meaning that there exist $a \in \mathbf{R}^n$, $a \neq 0$, such that $a^T(x - x_0) \leq 0$ for every $x \in \mathcal{C}$.*

3.2 Convex Functions

3.2.1 Domain of a function

The *domain* of a function $f : \mathbf{R}^n \to \mathbf{R}$ is the set $\mathbf{dom} f \subseteq \mathbf{R}^n$ over which f is well-defined, in other words:

$$\mathbf{dom}\, f := \{x \in \mathbf{R}^n : -\infty < f(x) < +\infty\}.$$

Here are some examples:

- The function with values $f(x) = \log(x)$ has domain $\mathbf{dom}\, f = \mathbf{R}_{++}$.
- The function with values $f(X) = \log\det(X)$ has domain $\mathbf{dom}\, f = \mathcal{S}^n_{++}$ (the set of positive-definite matrices).

3.2.2 Definition of convexity

A function $f : \mathbf{R}^n \to \mathbf{R}$ is convex if

i) $\mathbf{dom}\, f$ is convex;

ii) $\forall\, x, y \in \mathbf{dom}\, f$ and $\forall\, \lambda \in [0,1]$, $f\left(\lambda x + (1-\lambda)y\right) \leq \lambda f(x) + (1-\lambda) f(y)$.

Note that the convexity of the domain is required. For example, the function $f : \mathbf{R} \to \mathbf{R}$ defined as

$$f(x) = \begin{cases} x & \text{if } x \notin [-1, 1] \\ +\infty & \text{otherwise} \end{cases}$$

is *not* convex, although is it linear (hence, convex) on its domain $]-\infty, -1) \cup (1, +\infty[$.

We say that a function is *concave* if $-f$ is convex.

Here are some examples:

- The support function of any set is convex.
- The indicator function of a set is convex if and only if the set is convex.
- The quadratic function $f(x) = x^T P x + 2q^T x + r$, with $P \in \mathcal{S}^n_{++}$, is convex. (For a proof, see later.)
- The function $f : \mathbf{R} \to \mathbf{R}$ defined as $f(x) = 1/x$ for $x > 0$ and $f(x) = +\infty$ is convex.

3.2.3 Alternate characterizations of convexity

Let $f : \mathbf{R}^n \to \mathbf{R}$. The following are equivalent conditions for f to be convex.

- *Epigraph condition:* f is convex if and only if its *epigraph*

$$\mathbf{epi} f := \left\{(x,t) \in \mathbf{R}^{n+1} \ : \ t \geq f(x)\right\}$$

  is convex. We can us this result to prove for example, that the largest eigenvalue function $\lambda_{\max} : \mathcal{S}^n \to \mathbf{R}$, which to a given $n \times n$ symmetric matrix X associates its largest eigenvalue, is convex, since the condition $\lambda_{\max}(X) \leq t$ is equivalent to the condition that $tI - X \in \mathcal{S}^n_+$.

- *Restriction to a line:* The function f is convex if and only if its restriction to *any* line is convex, meaning that for every $x_0 \in \mathbf{R}^n$, and $v \in \mathbf{R}^n$, the function $g(t) := f(x_0 + tv)$ is convex.

  For example, the function $f(X) = \log\det X$ is convex. (Prove this as an exercise.) You can also use this to prove that the quadratic function $f(x) = x^T P x + 2q^T x + r$ is convex if and only if $P \succeq 0$.

- *First-order condition:* If f is differentiable (that is, $\mathbf{dom} f$ is open and the gradient exists everywhere on the domain), then f is convex if and only if

$$\forall\, x, y \ : \ f(y) \geq f(x) + \nabla f(x)^T (y - x).$$

  The geometric interpretation is that the graph of f is bounded below everywhere by anyone of its tangents.

- *Second-order condition:* If f is twice differentiable, then it is convex if and only if its Hessian $\nabla^2 f$ is positive semi-definite everywhere. This is perhaps the most commonly known characterization of convexity.

  For example, the function $f(x,t) = x^T x/t$ with domain $\{(x,t) \ : \ t > 0\}$, is convex. (Check this!) Other examples include the log-sum-exp function, $f(x) = \log \sum_{i=1}^n \exp x_i$, and the quadratic function alluded to above.

3.2.4 Operations that preserve convexity

- The nonnegative weighted sum of convex functions is convex.

- The composition with an affine function preserves convexity: if $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$ and $f : \mathbf{R}^m \to \mathbf{R}$ is convex, then the function $g : \mathbf{R}^n \to \mathbf{R}$ with values $g(x) = f(Ax + b)$ is convex.

- The *pointwise maximum* of a family of convex functions is convex: if $(f_\alpha)_{\alpha \in \mathcal{A}}$ is a family of convex functions index by α, then the function

$$f(x) := \max_{\alpha \in \mathcal{A}} f_\alpha(x)$$

  is convex. For example, the dual norm

$$x \to \max_{y \,:\, \|y\| \leq 1} y^T x$$

  is convex, as the maximum of convex (in fact, linear) functions (indexed by the vector y). Another example is the largest singular value of a matrix A: $f(A) = \sigma_{\max}(A) = \max_{x \,:\, \|x\|_2 = 1} \|Ax\|_2$. Here, each function (indexed by $x \in \mathbf{R}^n$) $A \to \|Ax\|_2$ is convex, since it is the composition of the Euclidean norm (a convex function) with an affine function $A \to Ax$.

Also, this can be used to prove convexity of the function we introduced in lecture 2,

$$\|x\|_{1,k} := \sum_{i=1}^{k} |x|_{[i]} = \max_{u} \; u^T|x| \;\; : \;\; \sum_{i=1}^{n} u_i = k, \;\; u \in \{0,1\}^n,$$

where we use the fact that for any u feasible for the maximization problem, the function $x \to u^T|x|$ is convex (since $u \geq 0$).

- If f is a convex function in $x = (y, z)$, then the function $g(y) := \min_z f(y, z)$ is convex. (Note that joint convexity in (y, z) is essential.)
- If f is convex, its *perspective* $g(x, t) := tf(x/t)$ with domain $\mathbf{dom}g = \{(x, t) : x \in \mathbf{dom}f, t > 0\}$, is convex. You can use this to prove convexity of the function $f(x, t) = x^T x/t$, with domain $\{(x, t) : t > 0\}$.
- The composition with another function does not always preserve convexity. However, if the functions $g_i : \mathbf{R}^n \to \mathbf{R}$, $i = 1, \ldots, k$ are convex and $h : \mathbf{R}^k \to \mathbf{R}$ is convex and non-decreasing in each argument, with $\mathbf{dom}g_i = \mathbf{dom}h = \mathbf{R}$, then $x \to (h \circ g)(x) := h(g_1(x), \ldots, g_k(x))$ is convex.

  For example, if g_i's are convex, then $\log \sum_i \exp g_i$ also is.

3.2.5 Conjugate function

The conjugate function of a function $f : \mathbf{R}^n \to \mathbf{R}$ is the function defined as

$$f^*(y) = \max_{x} \; x^T y - f(x) \;\; : \;\; x \in \mathbf{dom}f.$$

The function f^* is convex (even if f is not). The conjugate function plays a very important role in convex optimization, similar to the Fourier transform in signal processing.

For example, the conjugate of the convex quadratic function $f(x) = (1/2)x^T Q x$, with $Q \succ 0$, is $f^*(y) = (1/2)y^T Q^{-1} y$. Another important example is the conjugate of a norm, which is the indicator function of the unit ball for the dual norm:

$$f^*(y) = \begin{cases} 0 & \text{if } \|y\|_* \leq 1, \\ +\infty & \text{otherwise} \end{cases}$$

The conjugate of a conjugate is not always the original function. However, if f is convex, and closed (meaning that its epigraph is), then $f^{**} = f$.

3.3 Convex Optimization Problems

3.3.1 Terminology

Standard form. The problem

$$\begin{array}{lll} \min\limits_{x} \; f_0(x) \;\; : & f_i(x) \leq 0, & i = 1, \cdots, m, \\ & Ax = b, & i = 1, \cdots, p, \end{array} \tag{3.2}$$

is called a *convex optimization problem* if the objective function f_0 is convex; the functions defining the inequality constraints f_i, $i = 1, \ldots, m$ are convex; and $A \in \mathbf{R}^{p \times n}$, $b \in \mathbf{R}^p$ define the affine equality constraints. Note that, in the convex optimization model, we do not tolerate equality constraints other than affine ones.

Optimal value and feasible set. We usually denote by p^* the *optimal value* of the problem, and by $\mathcal{X}$ the *feasible set:*

$$\mathcal{X} = \{x \in \mathbf{R}^n \ : \ f_i(x) \leq 0, \quad i = 1, \cdots, m, \ \ Ax = b\}\,.$$

If $\mathcal{X}$ is empty, then we say that the problem is not feasible. By convention, in this case we set $p^* = +\infty$. The optimal value can also assume the value $-\infty$, in which case we say that the problem is *unbounded below.* An example of a problem that is unbounded below is an unconstrained problem with $f_0(x) = -\log x$, with domain $\mathbf{R}_{++}$.

Feasibility problems. In some instances, we do not care about any objective function, and simply seek a feasible point. This so-called *feasibility problem* can be formulated in the standard form, using a zero (or constant) objective.

3.3.2 Optimality

Local and global optima. A feasible point $x^* \in \mathcal{X}$ is a globally optimal (optimal for short) if $f_0(x) = p^*$.

A feasible point $x^* \in \mathcal{X}$ is a locally optimal if there exist $R > 0$ such that $f(x^*)$ equals the optimal value of problem (3.2) with the added constraint $\|x - x^*\| \leq R$. That is, x^* solves the problem "locally".

For convex problems, any locally optimal point is globally optimal.

Indeed, let x^* be a local minimizer of f_0 on the set $\mathcal{X}$, and let $y \in \mathcal{X}$. By definition, $x^* \in \mathbf{dom}\, f_0$. We need to prove that $f_0(y) \geq f_0(x^*) = p^*$. There is nothing to prove if $f_0(y) = +\infty$, so let us assume that $y \in \mathbf{dom}\, f_0$. By convexity of f_0 and $\mathcal{X}$, we have $x_\theta := \theta y + (1-\theta)x^* \in \mathcal{X}$, and:

$$f_0(x_\theta) - f_0(x^*) \leq \theta(f_0(y) - f_0(x^*)).$$

Since x^* is a local minimizer, the left-hand side in this inequality is nonnegative for all small enough values of $\theta > 0$. We conclude that the right hand side is nonnegative, i.e., $f_0(y) \geq f_0(x^*)$, as claimed.

Optimal set. The optimal set, $\mathcal{X}^{\rm opt}$, is the set of optimal points. This set may be empty: for example, the feasible set may be empty. Another example is when the optimal value is only reached in the limit; think for example of the case when $n = 1$, $f_0(x) = \exp x$, and there are no constraints.

In any case, the optimal set is convex, since it can be written

$$\mathcal{X}^{\rm opt} = \{x \in \mathbf{R}^n \ : \ f_0(x) \leq p^*, \ \ x \in \mathcal{X}\}\,.$$

Optimality condition. When f_0 is differentiable, then we know that for every $x, y \in \mathbf{dom}\, f_0$,

$$f_0(y) \geq f_0(x) + \nabla f_0(x)^T (y - x).$$

Then x is optimal if and only if

$$\forall\, y \in \mathcal{X} \ : \ \nabla f_0(x)^T (y - x) \geq 0.$$

If $\nabla f_0(x) \neq 0$, then it defines a supporting hyperplane to the feasible set at x.

Some examples of optimality conditions:

- For unconstrained problems, the optimality condition reduces to $\nabla f_0(x) = 0$.

- For problems with equality constraints only, the condition is that there exists a vector $\nu \in \mathbf{R}^p$ such that

$$x \in \mathbf{dom}\, f_0, \quad Ax = b, \quad \nabla f_0(x) = A^T \nu.$$

  Indeed, the optimality condition can be written as: $\nabla f_0(x)^T u \geq 0$ for every $u \in \mathcal{N}(A)$, which is the same as $\nabla f_0(x)^T u = 0$ for every $u \in \mathcal{N}(A)$. In turn, the latter means that $\nabla f_0(x)$ belongs to $\mathcal{R}(A^T)$, as claimed.

- For problems involving a single norm bound:

$$\min_x \; f_0(x) \; : \; \|x\| \leq 1,$$

  the condition reads

$$x \in \mathbf{dom}\, f_0, \quad \|x\| \leq 1, \quad -\nabla f_0(x)^T x \geq \|\nabla f_0(x)\|_*.$$

  From this, we conclude that if the constraint is not satisfied with equality at optimum, that is, $\|x\| < 1$, then $\nabla f_0(x) = 0$, and the problem is effectively unconstrained (it has the same solution as the unconstrained problem).

The optimality conditions given above might be hard to solve. We will return to this issue later.

3.3.3 Equivalent problems

We can transform a convex problem into an equivalent one via a number of transformations. Sometimes the transformation is useful to obtain an explicit solution, or is done for algorithmic purposes. The transformation does not necessarily preserve the convexity properties of the problem. Here is a list of transformations that do preserve convexity.

Epigraph form. Sometimes it is convenient to work with the equivalent *epigraph* form:

$$\min_{(x,t)} \; t \; : \; t \geq f_0(x), \quad x \in \mathcal{X},$$

in which we observe that we can always assume the cost function to be differentiable (in fact, linear), at the cost of adding one scalar variable.

Implicit constraints. Even though some problems appear to be unconstrained, they might contain implicit constraints. Precisely, the problem above has an implicit constraint $x \in \mathcal{D}$, where $\mathcal{D}$ is the problem's *domain*

$$\mathcal{D} := \mathbf{dom}\, f_0 \bigcap_{i=1}^m \mathbf{dom}\, f_i.$$

For example, the problem

$$\min_x \; c^T x - \sum_{i=1}^m \log(b_i - a_i^T x),$$

where $c \in \mathbf{R}^n$, $b \in \mathbf{R}^m$ and a_i^T's are the rows of $A \in \mathbf{R}^{m \times n}$, arises as an important sub-problem in some linear optimization algorithms. This problem has the implicit constraint that x should belong to the interior of the polyhedron $\mathcal{P} = \{x \; : \; Ax \leq b\}$.

Making explicit constraints implicit. The problem in standard form can be also written in a form that makes the constraints that are explicit in the original problem, implicit. Indeed, an equivalent formulation is the unconstrained convex problem

$$\min_x \ f(x)$$

where f is the sum of the original objective and the indicator function of the feasible set $\mathcal{X}$:

$$f(x) = f_0(x) + \mathbf{1}_{\mathcal{X}}(x) = \begin{cases} f_0(x) & x \in \mathcal{X} \\ +\infty & x \notin \mathcal{X}. \end{cases}$$

In the unconstrained above, the constraints are implicit. One of the main differences with the original, constrained problem is that now the objective function may not be differentiable, even if all the functions f_i's are.

A less radical approach involves the convex problem wih *one* inequality constraint

$$\min_x \ f_0(x) \ : \ Ax = b, \ \ g(x) := \max_{1 \le i \le m} \ f_i(x) \le 0,$$

which is equivalent to the original problem. In the above formulation, the structure of the inequality constraint is made implicit. Here, the reduction to a single constraint has a cost, since the function g may not be differentiable, even though all the f_i's are.

The above transformations show the versatility of the convex optimization model. They are also useful in the analysis of such problems.

Equality constraint elimination. We can eliminate the equality constraint $Ax = b$, by writing them as $x = x_0 + Nz$, with x_0 a particular solution to the equality constraint, and the columns of N span the nullspace of A. Then we can rewrite the problem as one without equality constraints:

$$\min_z \ f_0(Nz + x_0) \ : \ f_i(Nz + x_0) \le 0, \ \ i = 1, \ldots, m.$$

This transformation preserves convexity of the the function involved. In practice, it may not be a good idea to perform this elimination. For example, if A is sparse, the original problem has a sparse structure that may be exploited by some algorithms. In contrast, the reduced problem above does not inherit the sparsity characteristics, since in general the matrix N is dense.

Introducing equality constraints. We can also introduce equality constraints in the problem. There might be several justifications for doing so: to reduce a given problem to a standard form used by off-the-shelf algorithms, or to use in decomposition methods for large-scale optimization.

The following example shows that introducing equality constraint may allow to exploit sparsity patterns inherent to the problem. Consider

$$\min_{(x_k)_{k=1}^K, y} \ \sum_{k=1}^K f_{0,k}(x_k) \ : \ f_k(x_k, y) \le 0, \ \ k = 1, \ldots, K.$$

In the above the objective involves different optimization variables, which are coupled via the presence of the "coupling" variable y in each constraint. We can introduce K variables and rewrite the problem as

$$\min_{(x_k)_{k=1}^K, (y_k)_{k=1}^K, y} \ \sum_{k=1}^K f_{0,k}(x_k) \ : \ f_k(x_k, y_k) \le 0, \ \ y = y_k, \ \ k = 1, \ldots, K.$$

Now the objective and inequality constraints are all independent (they involve different optimization variables). The only coupling constraint is now an equality constraint. This can be exploited in parallel algorithms for large-scale optimization.

Slack variables. Sometimes it is useful to introduce slack variables. For example, the problem with affine inequalities

$$\min_x \ f_0(x) \ : \ Ax \leq b$$

can be written

$$\min_{x,s} \ f_0(x) \ : \ Ax + s = b, \ \ s \geq 0.$$

Minimizing over some variables. We may "eliminate" some variables of the problem and reduce it to one with fewer variables. This operation preserves convexity. Specifically, if f is a convex function of the variable $x \in \mathbf{R}^n$, and x is partitioned as $x = (x_1, x_2)$, with $x_i \in \mathbf{R}^{n_i}$, $i = 1, 2$, $n = n_1 + n_2$, then the function

$$\tilde{f}(x_1) := \min_{x=(x_1,x_2)} f(x_1, x_2)$$

is convex (look at the epigraph of this function). Hence the problem of minimizing f can be reduced to one involving x_1 only:

$$\min_{x_1} \ \tilde{f}(x_1).$$

The reduction may not be easy to carry out explicitly.

Here is an example where it is: consider the problem

$$\min_{x=(x_1,x_2)} \ x^T Q x \ : \ Ax_1 \leq b$$

where

$$Q := \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{pmatrix}$$

is positive semi-definite, Q_{22} is positive-definite, and $Q_{ij} \in \mathbf{R}^{n_i \times n_j}$, $i, j = 1, 2$. Since $Q \succ 0$, the above problem is convex. Furthermore, since the problem has no constraints on x_2, it is possible to solve for the minimization with respect to x_2 analytically. We end up with

$$\min_{x_1} \ x_1^T \tilde{Q} x_1 \ : \ Ax_1 \leq b$$

with $\tilde{Q} := Q_{11} - Q_{12} Q_{22}^{-1} Q_{12}^T$. From the reasoning above, we infer that $\tilde{Q}$ is positive semi-definite, since the objective function of the reduced problem is convex.

3.3.4 Maximization of convex functions

Sometimes we would like to *maximize* a convex function over a set $\mathcal{S}$. Such problems are usually hard to solve. The problem of maximizing the distance from a given point (say, 0) to a point in a polyhedron described as $\{x : Ax \leq b\}$ is an example of such hard problems.

One important property of convex functions is that their maximum over any set is the same as the maximum over the convex hull of that set. That is, for any set $\mathcal{S} \subseteq \mathbf{R}^n$ and any convex function $f : \mathbf{R}^n \to \mathbf{R}$, we have

$$\max_{x \in \mathcal{S}} f(x) = \max_{x \in \mathbf{Co}\mathcal{S}} f(x).$$

To prove this result is simple and I recommend you try.

In the example mentioned above, where we seek to maximize the Euclidean norm of a point in a polyhedron, if we know the vertices of the polyhedron, that is, we can express the polyhedron as $\mathcal{P} = \mathbf{Co}\{x^{(1)}, \ldots, x^{(K)}\}$, then the optimal distance is simply $\max_{1 \leq i \leq K} \|x^{(i)}\|_2$. Unfortunately, in practice, finding the vertices (given the original representation of $\mathcal{P}$ as an intersection of hyperplanes) is hard, and might involve an exponential number of vertices.

3.4 Linear Optimization

3.4.1 Definition and standard forms

Definition. A linear optimization problem (or, linear program, LP) is one of the standard form:

$$\min_x \ f_0(x) \ : \quad \begin{array}{l} f_i(x) \leq 0, \quad i = 1, \cdots, m, \\ Ax = b, \quad i = 1, \cdots, p, \end{array}$$

where every function $f_0, f_1, \ldots, f_m$ is affine. Thus, the feasible set of an LP is a polyhedron.

Standard forms. Linear optimization problems admits several standard forms. One is derived from the general standard form:

$$\min_x \ c^T x + d \ : \ Ax = b, \ \ Cx \leq h,$$

where the inequalities are understood componentwise[1]. The constant term d in the objective function is, of course, immaterial.

Another standard form—used in several off-the-shelf algorithms—-is

$$\min_x \ c^T x \ : \ Ax = b, \ \ x \geq 0.$$

We can always transform the above problem into the previous standard form, and vice-versa.

3.4.2 Examples

Piece-wise linear minimization. A piece-wise linear function is the pointwise maximum of affine functions, and has the form

$$f(x) = \max_{1 \leq i \leq m} \ (a_i^T x + b_i),$$

for appropriate vectors a_i and scalars b_i, $i = 1, \ldots, m$. The (unconstrained) problem of minimizing the piece-wise linear function above is not an LP. However, its epigraph form:

$$\min_{x,t} \ t \ : \ t \geq a_i^T x + b_i, \ \ i = 1, \ldots, m$$

is.

[1] Notice that the convention for componentwise inequalities differs from the one adopted in BV. I will reserve the symbol $\preceq$ or $\succeq$ for negative and positive semi-definiteness of symmetric matrices.

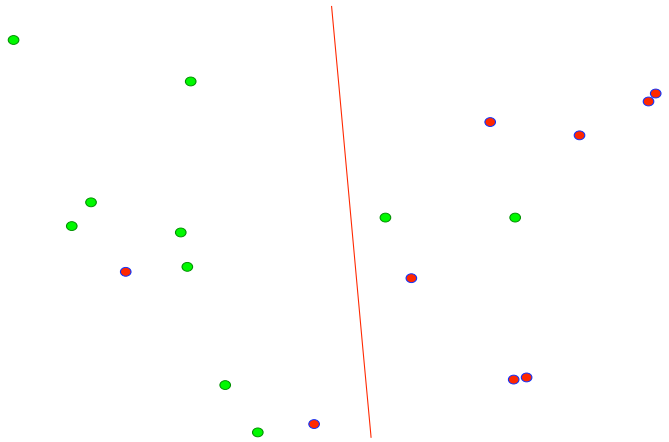

Figure 3.1: A linear classifier, with a total of four errors on the training set. The sum of the lengths of the dotted lines (which correspond to classification errors) is the value of the loss function at optimum.

l_1-norm regression. A related example involves the minimization of the l_1-norm of a vector that depends affinely on the variable. This arises in regression problems, such as image reconstruction. Using a notation similar to the previous example, the problem has the form

$$\min_x \sum_{i=1}^m |a_i^T x + b_i|.$$

The problem is not an LP, but we can introduce slack variables and re-write the above in the equivalent, LP format:

$$\min_{x,v} \sum_{i=1}^m v_i \ : \ -v_i \leq a_i^T x + b_i \leq v_i, \ \ i = 1, \ldots, m.$$

Linear binary classification. Consider a two-class classification problem as shown in Figure 3.1. Given m data points $x_i \in \mathbf{R}^n$, each of which is associated with a label $y_i \in \{-1, 1\}$, the problem is to find a hyperplane that separates, as much as possible, the two classes.

The two classes are separable by a hyperplane $\mathcal{H}(w, b) = \{x : w^T x + b \leq 0\}$, where $w \in \mathbf{R}^n$, $w \neq 0$, and $b \in \mathbf{R}$, if and only if $w^T x_i + b \geq 0$ for $y_i = +1$, and $w^T x_i + b \leq 0$ if $y_i = -1$. Thus, the conditions on (w, b)

$$y_i(w^T x_i + b) \geq 0, \ \ i = 1, \ldots, m$$

ensure that the data set is separable by a linear classifier. In this case, the parameters w, b allow to predict the label associated to a new point x, via $y = \mathbf{sign}(w^T x + b)$. The feasibility problem—finding (w, b) that satisfy the above separability constraints—is an LP. If the data set is strictly separable (every condition in (3.3) holds strictly), then we can re-scale the constraints and transform them into

$$y_i(w^T x_i + b) \geq 1, \ \ i = 1, \ldots, m. \tag{3.3}$$

In practice, the two classes may not be linearly separable. In this case, we would like to minimize, by proper choice of the hyperplane parameters, the total

number of classification errors. Our objective function has the form

$$\sum_{i=1}^{m} \psi(y_i(w^T x_i + b)),$$

where $\psi(t) = 1$ if $t < 0$, and 0 otherwise.

Unfortunately, the above objective function is not convex, and hard to minimize. We can replace it by an upper bound, which is called the *hinge* function, $h(t) = (1-t)_+ = \max(0, 1-t)$. Our problem becomes one of minimizing a piece-wise linear "loss" function:

$$\min_{w,b} \sum_{i=1}^{m} (1 - y_i(w^T x_i + b))_+.$$

In the above form, the problem is not yet an LP. We may introduce slack variables to obtain the LP form:

$$\min_{w,b,v} \sum_{i=1}^{m} v_i \ : \ v \geq 0, \ \ y_i(w^T x_i + b) \geq 1 - v_i, \ \ i = 1, \ldots, m.$$

The above can be seen as a variant of the separability conditions (3.3), where we allow for *infeasibilities*, and seek to minimize their sum. The value of the loss function at optimum can be read from Figure 3.1: it is the sum of the lengths of the dotted lines, from data points that are wrongly classified, to the hyperplane.

Network flow. Consider a network (directed graph) having m nodes connected by n directed arcs (ordered pairs (i, j)). We assume there is at most one arc from node i to node j, and no self-loops. We define the *arc-node incidence matrix* $A \in \mathbf{R}^{m \times n}$ to be the matrix with coefficients $A_{ij} = 1$ if arc j starts at node i, -1 if it ends there, and 0 otherwise. Note that the column sums of A are zero: $\mathbf{1}^T A = 0$.

A flow (of traffic, information, charge) is represented by a vector $x \in \mathbf{R}^n$, and the *total flow leaving node* i is then $(Ax)_i = \sum_{j=1}^{n} A_{ij} x_j$.

The *minimum cost network flow* problem has the LP form

$$\min_{x} \ c^T x \ : \ Ax = b, \ \ l \leq x \leq u,$$

where c_i is the cost of flow through arc i, l, u provide upper and lower bounds on x and $b \in \mathbf{R}^m$ is an *external supply* vector. This vector may have positive or negative components, as it represents supply and demand. We assume that $\mathbf{1}^T b = 0$, so that the total supply equals the total demand. The constraint $Ax = b$ represents the balance equations of the network.

A more specific example is the *max flow* problem, where we seek to maximize the flow between node 1 (the source) and node m (the sink). It bears the form

$$\min_{x,t} \ t \ : \ Ax = te, \ \ l \leq x \leq u,$$

with $e = (1, 0, \ldots, 0, -1)$.

LP relaxation of boolean problems. A boolean optimization problem is one where the variables are constrained to be boolean. An example of boolean problem is the so-called boolean LP

$$p^* = \min_{x} \ c^T x \ : \ Ax \leq b, \ \ x \in \{0, 1\}^n.$$

Such problems are non-convex, and usually hard to solve. The LP relaxation takes the form

$$p^*_{\rm LP} := \min_x \ c^T x \ : \ Ax \leq b, \ \ 0 \leq x \leq \mathbf{1}.$$

The relaxation provides a lower bound on the original problem: $p^*_{\rm LP} \leq p^*$. Hence, its optimal points may not be feasible (not boolean). Even though a solution of the LP relaxation may not necessarily be boolean, we can often interpret it as a *fractional* solution to the original problem. For example, in a graph coloring problem, the LP relaxation colors the nodes of the graph not with a single color, but with many.

Boolean problems are not always hard to solve. Indeed, in some cases, one can show that the LP relaxation provides an exact solution to the boolean problem, as optimal points turn out to be boolean. A few examples in this category, involving network flow problems with boolean variables:

- The *weighted bipartite matching* problem is to match N people to N tasks, in a one-to-one fashion. The cost of matching person i to task j is A_{ij}. The problem reads

$$\begin{array}{lll} & x_{ij} \in \{0,1\} & \\ \min_x \sum_{i,j=1}^N A_{ij}x_{ij} \ : & \forall\, j, \ \sum_{i=1}^N x_{ij} = 1 & \text{(one person for each task)} \\ & \forall\, i, \ \sum_{j=1}^N x_{ij} = 1 & \text{(one task for each person)} \end{array}$$

- The *shortest path* problem has the form

$$\min_x \ \mathbf{1}^T x \ : \ Ax = (1, 0, \ldots, 0, -1), \ \ x \in \{0,1\}^n,$$

  where A stands for the incidence matrix of the network, and arcs with $x_i = 1$ form a shortest forward path between nodes 1 and m. As before the LP relaxation in this case is exact, in the sense that its solution is boolean. (The LP relaxation problem can be solved very efficiently with specialized algorithms.)

3.5 Overview of conic optimization

3.5.1 Conic optimization models.

The linear optimization model can be written in standard form as

$$\min_x \ c^T x \ : \ Ax = b, \ \ x \geq 0,$$

where we express the feasible set as the intersection of an affine subspace $\{x : Ax = b\}$, with the non-negative orthant, $\mathbf{R}^n_+$. One can think of the linear equality constraints, and the objective, as the part in the problem that involves the data (A, b, c), while the sign constraints describe its structure.

With the advent of provably polynomial-time methods for linear optimization in the late 70's, researchers tried to generalize the linear optimization model, in a way that retained the nice complexity properties of the linear model.

Early attempts at generalizing the above model focussed on allowing the linear map $x \to Ax$ to be nonlinear. Unfortunately, as soon as we introduce non-linearities in the equality constraints, the model becomes non-convex and potentially intractable numerically. Thus, modifying the linear equality constraints is probably not the right way to go.

Instead, one can try to modify the "structural" constraints $x \in \mathbf{R}^n_+$. If one replaces the non-negative orthant with another set $\mathcal{K}$, then we obtain a

generalization of linear optimization. Clearly, we need $\mathcal{K}$ to be a convex set, and we can further assume it to be a cone (if not, we can always introduce a new variable and a new equality constraint in order to satisfy this condition). Hence the motivation for the so-called *conic optimization* model:

$$\min_x \; c^T x \;:\; Ax = b, \;\; x \in \mathcal{K}, \tag{3.4}$$

where $\mathcal{K}$ is a given convex cone.

The issue becomes then of finding those convex cones $\mathcal{K}$ for which one can adapt the efficient methods invented for linear optimization, to the conic problem above. A nice theory due to Nesterov and Nemirovski, which they developed in the late 80's, allows to find a rich class of cones for which the corresponding conic optimization problem is numerically tractable. We refer to this class as *tractable conic optimization.*

3.5.2 Tractable conic optimization.

The cones that are "allowed" in tractable conic optimization are of three basic types, and include any combination (as detailed below) of these three basic types. The three basic cones are

- The non-negative orthant, $\mathbf{R}^n_+$. (Hence, conic optimization includes linear optimization as a special case.)
- The second-order cone, $\mathcal{Q}^n := \{(x,t) \in \mathbf{R}^n_+ \;:\; t \geq \|x\|_2\}$.
- The semi-definite cone, $\mathcal{S}^n_+ = \{X = X^T \succeq 0\}$.

A variation on the second-order cone, which is useful in applications, involves the *rotated second-order cone* $\mathcal{Q}^n_{\rm rot} := \{(x,y,z) \in \mathbf{R}^{n+2} : 2yz \geq \|x\|_2^2, y \geq 0, z \geq 0\}$. We can easily convert the rotated second-order cone into the ordinary second-order cone representation, since the constraints $2yz \geq \|x\|_2^2$, $y \geq 0$, $z \geq 0$, are equivalent to

$$(y+z) \geq \left\| \begin{array}{c} (y-z) \\ \sqrt{2}\, x \end{array} \right\|_2 .$$

We can build all sorts of cones that are admissible for the tractable conic optimization model, using combinations of these cones. For example, in a specific instance of the problem, we might have constraints of the form

$$x_1 \geq 0, \;\; x_3 \geq \sqrt{x_1^2 + x_2^2}, \;\; \left(\begin{array}{cc} x_2 & x_4 \\ x_4 & x_5 \end{array} \right) \succeq 0.$$

The above set of constraints involves the non-negative orthant (first constraint), the second-order cone (second constraint), and the third, the semi-definite cone.

We can always introduce new variables and equality constraints, to make sure that the cone $\mathcal{K}$ is a direct product of the form $\mathcal{K}_1 \times \ldots \times \mathcal{K}_m$, where each $\mathcal{K}_i$ is a cone of one of the three basic types above. In the example above, since the variable x_2 appears in two of the cones, we add a new variable x_6 and the equality constraint $x_2 = x_6$. With that constraint, the constraint above can be written $x = (x_1, \ldots, x_6) \in \mathcal{K}$, where $\mathcal{K}$ is the direct product $\mathbf{R}_+ \times \mathcal{Q}^2 \times \mathcal{S}^2_+$.

Note that the three basic cones are nested, in the sense that we can interpret the non-negative orthant as the projection of a direct product of second-order cones on a subspace (think of imposing $x = 0$ in the definition of $\mathcal{Q}^n$). Likewise,

a projection of the semi-definite cone on a specific subspace gives the second-order cone, since

$$\|x\|_2 \leq t \Longleftrightarrow \begin{pmatrix} t & x_1 & \dots & x_n \\ x_1 & t & & 0 \\ \vdots & & \ddots & \\ x_n & 0 & & t \end{pmatrix} \succeq 0.$$

(The proof of this exercise hinges on the Schur complement lemma, see BV, pages 650-651.)

3.6 Second-order cone optimization

3.6.1 Definitions

3.6.2 Standard form.

We say that a problem is a second-order cone optimization problem (SOCP) if it is a tractable conic optimization problem of the form (3.4), where the cone $\mathcal{K}$ is a product of second-order cones and possibly the non-negative orthant $\mathbf{R}^n_+$.

A standard form for the SOCP model is

$$\min_x \ c^T x \ : \ Ax = b, \ \ \|C_i x + d_i\|_2 \leq e_i^T x + f_i, \ \ i = 1, \dots, m,$$

where we see that the variables $(C_i x + d_i, e_i^T x + f_i)$ should belong to a second-order cone of appropriate size. This corresponds to a convex problem in standard form, with the constraint functions $f_i(x) = \|C_i x + d_i\|_2 - (e_i^T x + f_i)$.

SOCPs contain LPs as special cases, as seen from the standard form above, with C_i, d_i all zero.

3.6.3 Special case: convex quadratic optimization.

Convex quadratic optimization (often written QP for short) corresponds to the convex optimization model

$$\min_x \ x^T Q x + c^T x \ : \ Cx \leq d, \ \ Ax = b,$$

where $Q = Q^T \succeq 0$. Thus, QPs are extensions of LPs where a convex, quadratic term is added to the linear objective.

We can view QPs as special cases of SOCP: first, we express the problem in a way to make the objective linear:

$$\min_{x,t} \ t + c^T x \ : \ Cx \leq d, \ \ Ax = b, \ \ t \geq x^T Q x,$$

then we observe that the last constraint can be expressed using a rotated second-order cone. Precisely, we have $t \geq x^T Q x$ if and only if $(Q^{1/2} x, t, 1) \in \mathcal{Q}^n_{\text{rot}}$.

3.6.4 Quadratically constrained, convex quadratic optimization.

QCQPs, as they are know by their acronym, correspond to problems of the form

$$\min_x \ q_0(x) \ : \ q_i(x) \leq 0, \ \ i = 1, \dots, m,$$

where the functions $q_0, \ldots, q_m$ are all convex and quadratic:

$$q_i(x) = x^T Q_i x + 2p_i^T x + r_i, \quad i = 1, \ldots, m,$$

with $Q_i \succeq 0$. Using rotated second-order cones, we can cast such problems as SOCPs.

Note that SOCPs cannot, in general, be cast as QCQPs. Consider a single SOC constraint of the form

$$\|Cx + d\|_2 \leq e^T x + f.$$

One may be tempted to square the SCO constraints and obtain a quadratic constraint of the form

$$\|Cx + d\|_2^2 \leq (e^T x + f)^2, \quad e^T x + f \geq 0.$$

While the above constraints are equivalent to the original SOC constraint, the first is not convex.

3.6.5 Examples

3.6.6 Risk-return trade-off in portfolio optimization

Consider the problem of investing in n assets, whose returns over one period (say, a month) are described as a random variable $y \in \mathbf{R}^n$, with mean $\hat{y}$ and covariance matrix Σ. A portfolio is described as a vector $x \in \mathbf{R}^n$, with x_i the amount invested in asset i (if no short-selling is allowed, we impose $x \geq 0$; in general, we might impose that the portfolio vector belongs to a given polyhedron $\mathcal{P}$). The return of the portfolio is then $x^T y$, and is a random variable with mean $x^T \hat{y}$ and variance $x^T \Sigma x$. The problem introduced by Markowitz seeks to find a trade-off between the expected return and the risk (variance) of the portfolio:

$$\max_{x \in \mathcal{P}} \ \hat{y}^T x - \gamma x^T \Sigma x,$$

where $\gamma > 0$ is a "risk-aversion" parameter. The above is a QP (convex quadratic program), a special case of SOCP.

3.6.7 Robust half-space constraint

Consider a constraint on $x \in \mathbf{R}^n$ of the form $a^T x \leq b$, with $a \in \mathbf{R}^n$ and $b \in \mathbf{R}$. Now assume that a is only known to belong to an ellipsoid $\mathcal{E} = \{\hat{a} + Ru \ : \ \|u\|_2 \leq 1\}$, with center $\hat{a} \in \mathbf{R}^n$ and $R \in \mathbf{R}^{n \times k}$ given. How can we guarantee that, irrespective of the choice of $a \in \mathcal{E}$, we still have $a^T x \leq b$?

The answer to this question hinges on the condition

$$b \geq \max_{a \in \mathcal{E}} \ a^T x = \hat{a}^T x + \max_{\|u\|_2 \leq 1} \ x^T R u = \hat{a}^T x + \|R^T x\|_2.$$

The above constraint is a second-order cone constraint.

3.6.8 Robust linear programming

Consider a linear optimization problem of the form

$$\min_x \ c^T x \ : \ a_i^T x \leq b_i, \quad i = 1, \ldots, m.$$

In practice, the coefficient vectors a_i may not be known perfectly, as they are subject to noise. Assume that we only know that $a_i \in \mathcal{E}_i$, where $\mathcal{E}_i$ are given

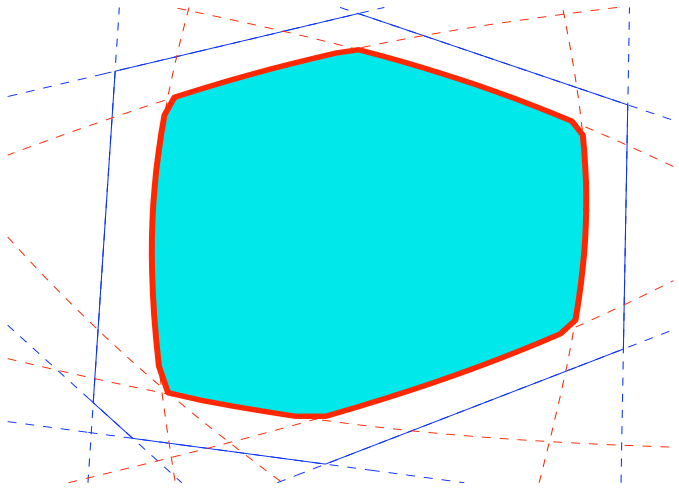

Figure 3.2: The robust feasible set associated to a linear optimization problem with row-wise spherical uncertainty on the coefficient matrix. The original feasible set is a polyhedron, with boundary shown in blue line. The robust feasible set is the intersection of robust half-space constraints, with boundaries shown as red dotted lines.

ellipsoids. In robust optimization, we seek to minimize the original objective, but we insist that each constraint be satisfied, irrespective of the choice of the corresponding vector $a_i \in \mathcal{E}_i$. Based on the earlier result, we obtain the second-order cone optimization problem

$$\min_x \; c^T x \;:\; \hat{a}_i^T x + \|R_i^T x\|_2 \leq b_i, \;\; i = 1, \ldots, m,$$

where $\mathcal{E}_i = \{\hat{a}_i + R_i u \;:\; \|u\|_2 \leq 1\}$. In the above, we observe that the feasible set is smaller than the original one, due to the terms involving the l_2-norms.

Figure (3.2) illustrates the kind of feasible set one obtains in a particular instance of the above problem, with spherical uncertainties (that is, all the ellipsoids are spheres, $R_i = \rho I$ for some $\rho > 0$). We observe that the robust feasible set is indeed contained in the original polyhedron.

3.6.9 Robust separation

In lecture 5, we have discussed the problem of linear separation of two classes of points. We revisit this example, assuming now that each point x_i, $i = 1, \ldots, m$, is only known to belong to an ellipsoid $\mathcal{E}_i = \{\hat{x}_i + R_i u \;:\; \|u\|_2 \leq 1\}$, where $\hat{x}_i$ is the "nominal" value, and matrix R_i describes the ellipsoidal uncertainty around it. The condition for an hyperplane $\mathcal{H}(w, b) = \{x \;:\; w^T x + b \leq 0\}$, where $w \in \mathbf{R}^n$, $w \neq 0$, and $b \in \mathbf{R}$, to separate the ellipsoids is

$$\forall\, i = 1, \ldots, m, \;\; \forall\, x \in \mathcal{E}_i \;:\; y_i(w^T x_i + b) \geq 0,$$

which is equivalent to the second-order cone constraints

$$w^T \hat{x}_i + b \geq \|R_i^T w\|_2, \;\; i = 1, \ldots, m.$$

Consider the special case when all the ellipsoids are spheres of given radius ρ, that is, $R_i = \rho I$, $i = 1, \ldots, m$. Now look for the hyperplane that is maximally

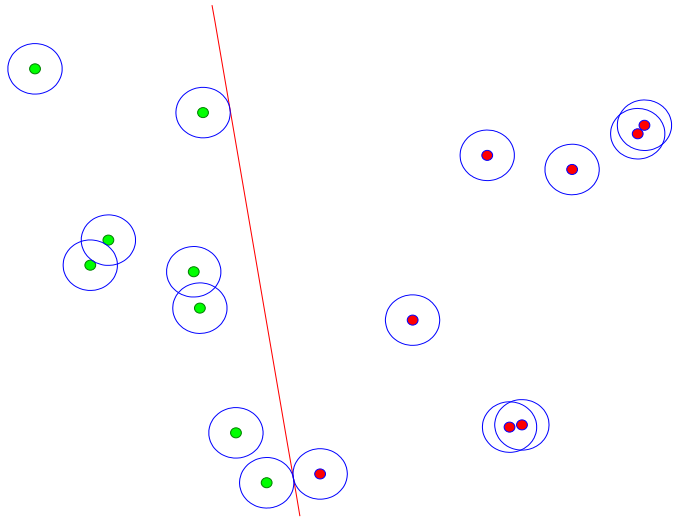

Figure 3.3: A linear classifier that separates data points with maximal spherical uncertainty around them.

robust, that is, tolerates the largest amount of uncertainty in the data points (as measured by ρ). Our problem writes

$$\max_{\rho,w,b} \rho \ : \ w^T \hat{x}_i + b \geq \rho \|w\|_2, \ \ i = 1, \ldots, m.$$

Since the constraints are homogeneous in (w, b), we can without loss of generality impose $\rho\|w\|_2 = 1$. We then formulate the above problem as the LP

$$\min_{w,b} \|w\|_2 \ : \ w^T \hat{x}_i + b \geq 1, \ \ i = 1, \ldots, m.$$

The quantity $\rho = 1/\|w^*\|_2$ is called the margin of the optimal classifier, and gives the largest amount of spherical uncertainty the data points can tolerate before they become not separable.

3.7 Semidefinite optimization

3.7.1 Definition and standard forms

Definition. We say that a problem is a semidefinite optimization problem (SDP) if it is a conic optimization problem of the form

$$\min_x c^T x \ : \ Ax = b, \ \ x \in \mathcal{K},$$

where the cone $\mathcal{K}$ is a product of semidefinite cones of given size. The decision variable x contains elements of matrices, and the constraint $x \in \mathcal{K}$ imposes positive semidefiniteness on such matrices.

Standard form. Often, it is useful to think of x as a *matrix* variable. The following standard form formalizes this.

First note that, without loss of generality, we may assume that $\mathcal{K} = \mathcal{S}^n_+$. Let us take an example with $\mathcal{K} = \mathcal{S}^{n_1}_+ \times \mathcal{S}^{n_2}_+$ to illustrate why. Indeed, the condition that $(X_1, X_2) \in \mathcal{S}^{n_1}_+ \times \mathcal{S}^{n_2}_+$ is the same as $X := \mathbf{diag}(X_1, X_2) \in$

$\mathcal{S}_+^{n_1+n_2}$. Thus, by imposing that certain elements of the matrix X be zero (those outside the diagonal blocks of size $n_i \times n_i$), we can always reduce our problem to the case when the cone $\mathcal{K}$ is the semidefinite cone itself, and the matrix variable is constrained to be block diagonal (with block sizes n_1, n_2), and live in that cone.

A standard form for the SDP model is thus

$$\min_{X \in \mathcal{S}^n} \mathbf{Tr}\, CX \;:\; \mathbf{Tr}(A_i X) = b_i, \;\; i = 1, \ldots, m, \;\; X \succeq 0,$$

where $C, A_1, \ldots, A_m \in \mathcal{S}^n$, and $b \in \mathbf{R}^m$. (In the above, we have used the fact that a single affine equality constraint on a symmetric matrix variable X can be represented as a scalar product condition of the form $\langle A, X \rangle = b$, for appropriate symmetric matrix A and scalar b.)

Inequality form. An alternate standard form, called the inequality standard form, is

$$\min_x c^T x \;:\; F_0 + \sum_{i=1}^n x_i F_i \succeq 0.$$

where matrices $F_0, \ldots, F_n \in \mathcal{S}^m$. The constraint in the above problem is called a *linear matrix inequality* (LMI). The above form is easily converted to the previous standard form (I suggest you try as an exercise).

3.7.2 Special cases

As discussed above, we can use the above formalism to handle multiple LMIs, using block-diagonal matrices. To illustrate this, we show that SDPs contain LPs as a special case. Indeed, the LP

$$\min_x c^T x \;:\; Ax \leq b$$

is equivalent to the SDP

$$\min_x c^T x \;:\; F(x) := \mathbf{diag}(b_1 - a_1^T x, \ldots, b_m - a_m^T x) \succeq 0,$$

where a_i^T is the i-th row of A. Thus, LPs are SDPs with a diagonal matrix in the LMIs.

Similarly, SDPs contain SOCPs, by virtue of the following result, already mentioned in lectue 5: for every x, t,

$$\|x\|_2 \leq t \Longleftrightarrow \begin{pmatrix} t & x_1 & \ldots & x_n \\ x_1 & t & & 0 \\ \vdots & & \ddots & \\ x_n & 0 & & t \end{pmatrix} \succeq 0.$$

Thus, a second-order cone constraint can be written as an LMI with an "arrow" matrix.

3.7.3 Examples

SDP relaxation of boolean problems. In lecture 5, we have seen LP relaxations for boolean LPs, which are LPs with the added non-convex requirement that the decision variable should be a boolean vector. This approach does not easily extend to the case when the problem involves quadratics.

To illustrate this, consider the *max-cut* problem. We are given a graph with vertices labeled $1, \ldots, n$, with weights $w_{ij} = w_{ji}$ for every pair of vertices (i, j). We seek a cut (a subset S of V) such that the total weight of all the edges that leave S is maximized. This can be expressed as the quadratic boolean problem

$$\max_x \ \frac{1}{2} \sum_{i<j} w_{ij}(1 - x_i x_j) \ : \ x \in \{-1, 1\}^n.$$

The problem is NP-hard. In 1995, Goemans and Williamson introduced an SDP relaxation (upper bound) of the problem, and showed that its value is at most within 15% of the optimal value of the combinatorial problem above, independent of problem size.

To explain the relaxation, let us embed the above problem into the more general problem of maximizing a given quadratic form over a boolean set:

$$p^* := \max_{x \in \{-1,1\}^n} \ x^T W x,$$

where $W \in \mathcal{S}^n$ is a given symmetric matrix. (As an exercise, try to cast the max-cut problem into the above formalism.)

We can express the problem as

$$p^* = \max_X \ \mathbf{Tr}\, WX \ : \ X \succeq 0, \ \ X_{ii} = 1, \ \ i = 1, \ldots, n, \ \ \mathbf{rank} X = 1.$$

Indeed, X is feasible for the above if and only $X = xx^T$ for some $x \in \{-1, 1\}^n$, in which case $\mathbf{Tr}\, WX = x^T W x$.

Relaxing (meaning: ignoring) the rank constraint leads to the upper bound $p^* \leq d^*$, where

$$d^* := \max_X \ \mathbf{Tr}\, WX \ : \ X \succeq 0, \ \ X_{ii} = 1, \ \ i = 1, \ldots, n.$$

The above is an SDP (in standard form). Nesterov has shown that, for arbitrary matrices W, the above relaxation is within $\pi/2$ of the true value, that is, $p^* \leq d^* \leq (\pi/2)p^*$.

Non-convex quadratic optimization. The approach used above can be applied to general non-convex quadratic optimization, which has the form

$$\min_x \ q_0(x) \ : \ q_i(x) \leq 0, \ \ i = 1, \ldots, m,$$

where $x \in \mathbf{R}^n$ is the decision variable, and q_i's are quadratic functions, of the form

$$q_i(x) := x^T Q_i x + 2 q_i x + p_i, \ \ i = 1, \ldots, m,$$

with $Q_i \in \mathcal{S}^n$, $q_i \in \mathbf{R}^n$ and $p_i \in \mathbf{R}$ given. The above problem is not convex in general (we have not imposed positive semi-definiteness on the Q_i's).

We note that $q_i(x) = L_i(xx^T, x)$, with $L_i : \mathcal{S}^n \times \mathbf{R}^n \to \mathbf{R}$ the affine functions:

$$L_i(X, x) := \mathbf{Tr}\, XQ_i + 2q_i x + p_i, \ \ i = 1, \ldots, m.$$

We can express the problem as

$$\min_{X,x} \ L_0(X, x) \ : \ L_i(X, x) \leq 0, \ \ i = 1, \ldots, m, \ \ X \succeq 0, \ \ \mathbf{rank}(X) = 1.$$

Dropping the rank constraint leads to an SDP relaxation (lower bound):

$$\min_{X,x} \ L_0(X, x) \ : \ L_i(X, x) \leq 0, \ \ i = 1, \ldots, m, \ \ X \succeq 0.$$

The above relaxation can be arbitrarily bad, but in some cases it is exact. For example, in the case of a *single* inequality constraint ($m = 1$), the SDP relaxation provides the exact value of the original non-convex problem.

Stability analysis of uncertain dynamical systems. Consider a time-varying dynamical system of the form

$$x(t+1) = A(t)x(t), \;\; t = 0, 1, 2, \ldots \tag{3.5}$$

with $x(t) \in \mathbf{R}^n$ the state vector, and $A(t) \in \mathbf{R}^{n\times n}$. We assume that all is known about $A(t)$ is that $A(t) \in \mathcal{A} = \{A_1, \ldots, A_L\}$, with $A_i \in \mathbf{R}^{n\times n}$, $i = 1, \ldots, L$ given matrices. We say that the system is asymptotically stable if, for every initial condition $x(0) \in \mathbf{R}^n$, the resulting trajectory goes to zero: $x(t) \to 0$ as $t \to +\infty$. Note that ascertaining asymptotic stability of the above "switched" system is hard in general, so we look for a tractable sufficient condition for asymptotic stability.

Let us denote by $\|\cdot\|$ the largest singular value norm. Clearly, if for every $t \geq 0$, we have $\|A(t)\| \leq \sigma < 1$ for some $\sigma < 1$, then $\|x(t+1)\|_2 \leq \sigma\|x(t)\|_2$, which shows that the state vector goes to zero at least as fast as σ^t. The norm condition "$\|A(t)\| < 1$ for every t" is conservative, since it implies that the state decreases monotonically.

To refine the norm-based sufficient condition for asymptotic stability, we can use scaling. For $S \in \mathbf{R}^{n\times n}$ a given invertible matrix, we define $\bar{x}(t) := Sx(t)$. In the new state space defined by $\bar{x}$, the state equations become $\bar{x}(t+1) = \bar{A}(t)\bar{x}(t)$, with $\bar{A}(t) := SA(t)S^{-1}$. The asymptotic stability of the system is independent of the choice of S, since $\bar{x}(t)$ converges to zero if and only if $x(t)$ does. However, the norm-based sufficient condition for asymptotic stability is not independent of the choice of S. Indeed, if we impose that condition on $\bar{A}(t)$, we obtain $\|SA(t)S^{-1}\| < 1$. In turn, this condition writes[2]

$$A(t)^T P A(t) \prec P,$$

where $P = S^T S \succ 0$. (The original norm-based sufficient condition is recovered with $P = I$.)

We conclude that the existence of $P \in \mathcal{S}^n$ such that

$$P \succ 0, \;\; \forall\, i = 1, \ldots, L \;:\; A_i^T P A_i \prec P$$

ensures the stability of the system (3.5), regardless of the choice of the trajectory of the matrix $A(t) \in \mathcal{A}$. The above condition is an LMI in matrix variable P.

Note that since for fixed $P \succ 0$, the set $\{A : A^T P A \prec P\}$ is convex, requiring that it contains the finite set $\mathcal{A}$ is the same as requiring that it contains its convex hull. So the above condition also ensures stability of the system (3.5), when $A(t)$ is allowed to be any time-varying convex combination of the matrices A_i, $i = 1, \ldots, L$.

Also, note that when $L = 1$, the above condition turns out to be exact, and equivalent to the condition that all the eigenvalues of the matrix A lie in the interior of the unit circle of the complex plane. This is the basic result of the so-called Lyapunov asymptotic stability theorem for linear systems.

3.8 More on non-convex quadratic optimization

3.8.1 More on rank relaxation.

Let us return to a problem involving (possibly non-convex) quadratic functions only. We consider a maximization problem instead of minimization, as can be done freely for non-convex problems:

$$p^* := \max_x \; q_0(x) \;:\; q_i(x) \leq 0, \;\; i = 1, \ldots, m, \tag{3.6}$$

[2]Here, we use the fact that for a matrix A, the condition $\|A\| < 1$ is equivalent to $A^T A \prec I$ (try to show this).

where $x \in \mathbf{R}^n$ is the decision variable, and q_i's are quadratic functions, of the form

$$q_i(x) := x^T Q_i x + 2q_i x + p_i, \quad i = 1, \ldots, m,$$

We have seen a rank-relaxation approach, which leads to an SDP relaxation $p^* \leq d^*$, where

$$d^* := \max_{X,x} \; L_0(X,x) \; : \; L_i(X,x) \leq 0, \quad i = 1, \ldots, m, \quad X \succeq 0,$$

where $L_i(X,x) := \mathbf{Tr}\, XQ_i + 2q_i x + p_i$, $i = 1, \ldots, m$.

The rank relaxation for the general problem above can be arbitrarily poor in quality (that his, the lower bound obtained this way is arbitrarily smaller than the optimal value of the initial problem). However, when the constraints are convex ($Q_i \succeq 0$), and the q_i's are all zero, then a result by Nemirovski, Roos and Terlaky (1999) shows that

$$\frac{d^*}{2\log(2m\mu)} \leq p^* \leq d^*, \quad \mu := \min(m, \max_{1 \leq i \leq m} \mathbf{Rank}(A_i)).$$

The measure μ of the quality of the approximation grows slowly with problem size (and is independent of the number of variables). This generalizes the earlier result by Goemans and Williamson (1994) and Nesterov (1998), which are valid for special cases of the above problem. The bound above comes also with a method to provide a sub-optimal solution with value guaranteed to be in the interval $[\frac{d^*}{2\log(2m\mu)}, d^*]$.

3.8.2 Lagrange relaxation

There is an alternative approach to bounding the non-convex quadratic problem. This other approach is part of a very general class of relaxations, to be discussed later, called Lagrange relaxations. We will also see later that, for generic quadratic problems, the Lagrange relaxation and the rank relaxation are essentially the same. It is however instructive to look into the Lagrange relaxation approach now, as an independent approach.

The idea is that if, for a given m-vector $\lambda \geq 0$, and scalar t, we have

$$\forall\, x \; : \; q_0(x) \leq \sum_{i=1}^m \lambda_i q_i(x) + t,$$

then for every x that is feasible for (3.6), the sum in the above is non-positive. Hence, $q_0(x) \geq t$, so that t is an upper bound on our problem. The condition above is easy to check, as it involves a single quadratic function: indeed, it is equivalent to the LMI in (t, λ):

$$\left(\begin{array}{cc} Q_0 & q_0 \\ q_0^T & r_0 \end{array} \right) \preceq \sum_{i=1}^m \lambda_i \left(\begin{array}{cc} Q_i & q_i \\ q_i^T & r_i \end{array} \right) + \left(\begin{array}{cc} 0 & 0 \\ 0 & t \end{array} \right). \tag{3.7}$$

Hence, the best upper bound that we can achieve using this approach is the SDP

$$\min_{t,\lambda} \; t \; : \; (3.7), \quad \lambda \geq 0.$$

As noted before, the SDP above and the SDP obtained via rank-relaxation in fact provide the same bound (provided some technical condition holds). We will explore these ideas in more detail later.

3.8.3 The $\mathcal{S}$-lemma

This mysterious name corresponds to a special case of non-convex quadratic optimization, where there is only a *single* constraint. (Refer to appendix B of [BV] for more details.)

The problem bears the form

$$\max_x \; q_0(x) \;:\; q_1(x) \leq 0,$$

where both q_0, q_1 are arbitrary quadratic functions. The $\mathcal{S}$-lemma states that if there exist a point $x \in \mathbf{R}^n$ such that $q_1(x) < 0$, then the rank relaxation and the (equivalent) Lagrange relaxation are both exact. The latter has the form

$$\min_{t,\lambda} \; t \;:\; \begin{pmatrix} Q_0 & q_0 \\ q_0^T & r_0 \end{pmatrix} \preceq \lambda \begin{pmatrix} Q_1 & q_1 \\ q_1^T & r_1 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & t \end{pmatrix}, \;\; \lambda \geq 0. \tag{3.8}$$

This shows in particular that the apparently non-convex problem of finding the direction of maximal variance for a given covariance matrix Σ, is actually convex. We have

$$\max_{x \neq 0} \frac{x^T \Sigma x}{x^T x} = \max_{x \,:\, \|x\|_2 = 1} x^T \Sigma x = \max_{x \,:\, \|x\|_2 \leq 1} x^T \Sigma x,$$

where the last inequality exploits the convexity of the quadratic function $x \to x^T \Sigma x$. The rank relaxation is

$$\max_X \; \mathbf{Tr}\, \Sigma X \;:\; \mathbf{Tr}\, X = 1, \;\; X \succeq 0.$$

The Lagrange relaxation yields (check this!)

$$\min_t \; t \;:\; tI \succeq \Sigma.$$

As expected, both give the same bound, and that bound is exact.

The $\mathcal{S}$-lemma has many applications, and we'll visit some of them below.

3.9 Optimization over ellipsoids

There is a strong connection between positive-definite matrices and ellipsoids, in the sense that optimization problems involving ellipsoids often reduce to (or can be approximated by) semi-definite optimization problems. This is not surprising, since ellipsoids are set defined by a single convex quadratic constraint.

3.9.1 Parametrizations of ellipsoids

A non-degenerate ellipsoid can be expressed as

$$\mathcal{E} := \left\{ x \;:\; (x - \hat{x})^T P^{-1} (x - \hat{x}) \leq 1 \right\}, \tag{3.9}$$

where $P \succ 0$. An alternate parametrization is

$$\mathcal{E} := \left\{ x = \hat{x} + Ru \;:\; \|Ru\|_2 \leq 1 \right\}, \tag{3.10}$$

where $R := P^{1/2} \succ 0$. A third parametrization is

$$\mathcal{E} = \left\{ x \;:\; q(x) := x^T Q x + 2q^T x + r \leq 0 \right\}, \tag{3.11}$$

where $Q \succ 0$ and $r < q^T Q^{-1} q$ (otherwise, the ellipsoid has empty interior). We have the correspondence $P = Q^{-1}$, $q = -P^{-1}\hat{x}$.

The parameters of the above representation have geometric interpretations. Obviously, $\hat{x}$ is the center of the ellipsoid. Furthermore, the eigenvalue decomposition of $P = U^T \Lambda U$, with $\lambda = \mathbf{diag}(\lambda) \succ 0$ and U orthogonal, can be interpreted geometrically as follows. The transformation U, when applied to the points in the ellipsoid, rotates them so that the ellipsoid has axes parallel to the coordinate axes. Precisely, letting $\bar{x} = U(x - \hat{x})$ allows to write the constraint $(x-\hat{x})^T P^{-1}(x-\hat{x}) \leq 1$ as $\bar{x}^T \Lambda \bar{x} \leq 1$, which corresponds to an ellipsoid parallel to the axes, with λ_i the semi-axis lengths.

You may wonder why we bother about different ways to describe ellipsoids. The reason is precisely that some problems involving ellipsoids are convex when using one parametrization, but may not using the others.

Often a measure of size of the ellipsoid is needed. We can measure the size as the sum of the semi-axis lengths, that is $\mathbf{Tr}\, P$. Note that this measure can be used to maximize or minimize the size, since it is linear. In terms of the other two parametrizations, $\mathbf{Tr}\, P = \mathbf{Tr}\, Q^{-1} = \|R\|_F^2$ is convex in Q and R.

Alternatively, the volume can be used. The volume of the set described by (3.10) is $\det R$, so it can be maximized with the concave function $P \rightarrow \log\det P$. If one is interested in minimizing volume, then the last two parametrizations (3.10) or (3.11) can be used, as now the volume is proportional to $\det Q^{-1}$ or $\log\det R^{-1}$, and $Q \rightarrow -\log\det Q$ is convex.

3.9.2 Examples

Consider the problem of checking wether a given ellipsoid $\mathcal{E}_0$ contains another one, $\mathcal{E}_1$. We assume that both $\mathcal{E}_0, \mathcal{E}_1$ have non-empty interior. We use the parametrization (3.11), with subscripts on the parameters (Q, q, r). The problem can be formulated as checking if the optimal value of the problem

$$p^* := \max_{x \in \mathcal{E}_1} q_0(x) = \max_x \{q_0(x) \ : \ q_1(x) \leq 0\}$$

is less or equal to zero. Since $\mathcal{E}_1$ has non-empty interior, we can apply the $\mathcal{S}$-lemma. The condition (3.8), written with $t = 0$, yields the necessary and sufficient condition for $\mathcal{E}_1 \subseteq \mathcal{E}_0$:

$$\left(\begin{array}{cc} Q_0 & q_0 \\ q_0^T & r_0 \end{array} \right) \preceq \lambda \left(\begin{array}{cc} Q_1 & q_1 \\ q_1^T & r_1 \end{array} \right), \ \ \lambda \geq 0.$$

The above condition is easily shown to be sufficient: multiply both terms on the left by z^T and on the right by z, with $z = (x, 1)$.

3.9.3 Maximum volume ellipsoid in a polyhedron.

We are given a polyhedron described by a set of affine inequalities

$$\mathcal{P} = \{x \ : \ a_i^T x \leq b_i, \ \ i = 1, \ldots, m\}.$$

We seek an ellipsoid of minimum volume that contains $\mathcal{P}$. We use the parametrization (3.10) for a candidate ellipsoid $\mathcal{E}$, and formulate the condition $\mathcal{P} \subseteq \mathcal{E}$ as

$$\forall\, i, \ \ \forall x \ : \ q(x) \leq 1 \Longrightarrow a_i^T x \leq b_i.$$

We can use the $\mathcal{S}$-lemma to treat these conditions. An equivalent and more direct way is to express the above as

$$\forall\, i, \ \ b_i \geq \max_{\|u\|_2 \leq 1} a_i^T(\hat{x} + Ru) = a_i^T \hat{x} + \|Ra_i\|_2.$$

We obtain the problem in the form

$$\min_R \log\det R^{-1} \ : \ \|Ra_i\|_2 + a_i^T\hat{x} \le b_i, \ \ i = 1, \ldots, m.$$

The problem is not an SDP but is convex (and algorithms for SDPs can be adapted to solve it).

Minimum trace ellipsoid containing points. We are given m points in $\mathbf{R}^n$, $x_1, \ldots, x_m$. We seek to find a minimum-trace ellipsoid that contains all the points. Since an ellipsoid is convex, containing the points is the same as containing their convex hull.

Here, we use the parametrization (3.11). The condition for a candidate ellipsoid to contain the points is

$$q(x_i) = x_i^T Q x_i + 2q^T x_i + r \le 0, \ \ i = 1, \ldots, m.$$

The above constraints on (Q, q, r) are affine. Hence the problem reduces to

$$\min_Q \mathbf{Tr}\, Q^{-1} \ : \ Q \succeq 0, \ \ x_i^T Q x_i + 2q^T x_i + r \le 0, \ \ i = 1, \ldots, m.$$

The above problem is convex, but not quite an SDP yet. We can introduce a new matrix variable and use Schur complements, to express the above as

$$\min \mathbf{Tr}\, X \ : \ \begin{pmatrix} X & I \\ I & Q \end{pmatrix} \succeq 0, \ \ x_i^T Q x_i + 2q^T x_i + r \le 0, \ \ i = 1, \ldots, m.$$

Alternatively, the parametrization (3.9) can be used. One can also work with volume as a measure of size. I suggest you try these variations as an exercise.

3.9.4 Minimum trace ellipsoid containing the sum of ellipsoids.

This is a variation on the problem examined in [BV,page 411].

We are given ellipsoids $\mathcal{E}_i$, $i = 1, \ldots, m$, and seek a minimum-trace outer ellipsoidal approximation $\mathcal{E}_0$ to their sum, which we define as

$$\mathcal{C} := \mathcal{E}_1 \oplus \ldots \oplus \mathcal{E}_m = \left\{ \sum_{i=1}^m x_i \ : \ x_i \in \mathcal{E}_i, \ \ i = 1, \ldots, m \right\}.$$

To simplify, assume that all the ellipsoids involved have the same center at 0, and let us use parametrization (3.11), with $q_i = 0$, $r_i = -1$, $i = 0, \ldots, m$. The condition $\mathcal{C} \subseteq \mathcal{E}$ for a candidate zero-center ellipsoid $\mathcal{E}$ is equivalent to

$$\forall\, x = (x_1, \ldots, x_m) \ : \ q_0(\sum_{i=1}^m x_i) \le 0 \text{ whenever } q_i(x_i) \le 0, \ \ i = 1, \ldots, m.$$

We can apply Lagrange relaxation and obtain the sufficient condition: there exist $\lambda \ge 0$ such that

$$\forall\, x = (x_1, \ldots, x_m) \ : \ q_0(\sum_{i=1}^m x_i) \le \sum_{i=1}^m \lambda_i q_i(x_i).$$

With the notation $M := [I, \ldots, I]$ and $Q(\lambda) := \mathbf{diag}(\lambda_i Q_i)_{i=1}^m$, and $P_i = Q_i^{-1}$, $i = 1, \ldots, m$, we write the above in the equivalent form

$$M^T Q_0 M \preceq Q(\lambda).$$

We can reduce the size of the matrices involved in this condition drastically. Let us assume $\lambda > 0$. Using Schur complements, and with $P_0 = Q_0^{-1}$, we re-write the above as

$$\left(\begin{array}{cc} Q(\lambda) & M^T \\ M & P_0 \end{array} \right) \succeq 0.$$

Using Schur complements again (albeit in the other direction!) we obtain $P_0 \succeq MQ(\lambda)^{-1}M^T$. With $\tau_i = 1/\lambda_i$, we obtain

$$P_0 \succeq \sum_{i=1}^{m} \tau_i P_i.$$

A minimum trace ellipsoid can be obtained by the SDP

$$\min_{P_0,\tau} \mathbf{Tr}\, P_0 \ : \ P_0 \succeq \sum_{i=1}^{m} \tau_i P_i, \ \ \tau \geq 0.$$

3.9.5 Application: reachable sets for discrete-time dynamical systems

Consider the discrete-time linear system

$$x(t+1) = Ax(t) + Bp(t), \ \ t = 0, 1, 2, \ldots$$

where $A \in \mathbf{R}^{n\times n}$ and $B \in \mathbf{R}^{n\times n_p}$. We assume that the initial condition $x(0)$ is known, while the signal p is considered to be noise, and is only known to be norm-bounded, precisely $\|p(t)\|_2 \leq 1$ for ever $t \geq 0$. The goal of *reachability analysis* is to come up with bounds on the state at a certain time T.

We can seek a minimum-trace ellipsoid that is guaranteed to contain $x(T)$, irrespective of the values of the pertubation signal $p(t)$ within its bounds. By applying the recursion, we can express $x(T)$ as a linear combination of $p(0), \ldots, p(T-1)$:

$$x(T) = Ax(0) + \sum_{t=0}^{T-1} A^t B p(t).$$

The problem of finding the minimum trace ellipsoid that contains the state vector at time T is thus a direct application of the ellipsoid-sum problem discussed previously.

For example, when $x(0) = 0$, the center of $\mathcal{E}_0$ can be shown to be zero, so we set $\mathcal{E}_0 = \{x \ : \ x^T Q_0 x \leq 1\}$. With $L(t) := A^t B$, $L := [L(0), \ldots, L(T-1)]$, we obtain the sufficient condition for $\mathcal{E}_0$ to contain the state vector at time T:

$$\exists\, \lambda \geq 0 \ : \ \forall\, p = (p(0), \ldots, p(T-1)), \ \ p^T L^T Q_0 L p \leq 1 + \sum_{t=0}^{T-1} \lambda_t (p(t)^T p(t) - 1).$$

The above is equivalent to $D_\lambda \succeq L^T Q_0 L$, $\sum_{t=0}^{T-1} \lambda_t \leq 1$, where $D_\lambda = \mathbf{diag}(\lambda_0 I_{n_p}, \ldots, \lambda_{T-1} I_{n_p})$.

3.10 Geometric Programming

3.10.1 Standard form

Monomials and posynomials. A *monomial* is a function $f : \mathbf{R}^n_{++} \to \mathbf{R}$, with values $f(x) = c x_1^{a_1} \ldots x_n^{a_n}$, where $c > 0$ and $a \in \mathbf{R}^n$. A *posynomial* is a non-negative combination (sum) of monomials. For example, $3.4 x_1^{-0.3} x_2^4$ is a monomial, $x_1^{-2} + 3 x_2 x_4^{-\pi}$ is a posynomial, while $x_1^{-2} - 3 x_2 x_4^{-\pi}$ is neither.

Standard form of a GP. A geometric program (GP) is a problem of the form

$$\min_x \ f_0(x) \ : \ f_i(x) \leq 1, \ \ i = 1, \ldots, m, \ \ g_i(x) = 1, \ \ i = 1, \ldots, p$$

where $f_0, \ldots, f_m$ are posynomials, and $g_1, \ldots, g_p$ are monomials.

Convex representation. GPs are not convex, but a simple change of variables $y_i = \log x_i$ leads to an equivalent, convex formulation.

First consider a monomial equality constraint: $g(x) = 1$, with $g(x) = cx_1^{a_1} \ldots x_n^{a_n}$, $c > 0$, $a \in \mathbf{R}^n$. Taking logarithms, the constraint can be written as $a^T y = b := -\log c$. Hence, monomial (equality) constraints on x-variables translate into affine (equality) constraints on the y-variables.

For a posynomial

$$f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1,k}} \ldots x_n^{a_{n,k}}$$

the constraint $f(x) \leq 1$ takes the form

$$F(y) := \log(f(e^y)) = \log \left(\sum_{k=1}^{K} e^{a_k^T y - b_k} \right) \leq 0,$$

where $a_k = (a_{1,k}, \ldots, a_{n,k})$, $b_k = -\log c_k$, $k = 1, \ldots, K$. Since F is a convex function, the above constraint is convex.

We obtain that any GP can, after a logarithmic change of variables, be equivalently expressed as a convex optimization problem of the form

$$\min_y F_0(y) \ : \ F_i(y) \leq 0, \ \ i = 1, \ldots, m, \ \ Ay = b,$$

where the functions F_i are of the form above, and the equality constraints correspond to the monomial ones in the original problem.

3.10.2 Examples

A geometric problem. This example is taken from the aforementioned article. Here we optimize the shape of a box-shaped structure with height h, width w, and depth d. We have a limit on the total wall area $2(hw + hd)$, and a limit the floor area wd, as well as lower and upper bounds on the aspect ratios h/w and w/d. Subject to these constraints, we wish to maximize the volume of the structure, hwd. This leads to the problem

$$\max_{h,w,d} \ hwd \ : \ 2(hw + hd) \leq A_{\text{wall}}, \ \ wd \leq A_{\text{floor}}, \ \ \alpha \leq h/w \leq \beta, \ \ \gamma \leq d/w \leq \delta.$$

Here d, h, and w are the optimization variables, and the problem parameters are A_{wall} (the limit on wall area), Aflr (the limit on floor area), and α, β, γ, δ (the lower and upper limits on the wall and floor aspect ratios). This problem can be cast as a GP. I suggest you try to put it in standard form.

Minimizing the spectral radius of a non-negative matrix. Assume $A \in \mathbf{R}^{n \times n}$ has non-negative entries, and is irreducible (meaning, $(I + A)^{n-1}$ is component-wise positive). The *Perron-Frobenius theorem* states that A has a real, positive eigenvalue λ_{pf}, (not necessarily unique) such that $\lambda_{\text{pf}} \geq |\lambda_i|$ for all other (possibly complex) eigenvalues λ_i. Hence, λ_{pf} is equal to spectral radius

of the matrix A and is a measure of rate of growth or decay of A^k as $k \to \infty$. The value $\lambda_{\rm pf}$ is called the Perron-Frobenius eigenvalue of A.

The Perron-Frobenius eigenvalue can be found using an optimization problem as follows:

$$\lambda_{\rm pf}(A) = \min_{\lambda, v} \ \lambda \ : \ Av \leq \lambda v, \ \ v > 0$$

The component-wise inequality $Av \leq \lambda v$ can be expressed as a set of posynomial inequalities in the elements of A, v, and λ, as:

$$\sum_j A_{ij} v_j \leq \lambda v_i, \ \ i = 1, \ldots, n.$$

In many applications, the goal is to minimize $\lambda_{\rm pf}(A(x))$ where elements of $A(x)$ are given posynomials in decision variables $x > 0$. We can formulate this problem as a GP:

$$\begin{array}{ll} \min \ \lambda \ : & \displaystyle\sum_j \frac{(A(x))_{ij} v_j}{\lambda v_i} \leq 1, \ \ i = 1, \ldots, m, \\ & x > 0, \ \ v > 0 \end{array}$$

As a specific application, consider a model of population dynamics of bacteria. Let $p(t) = (p_1(t), \ldots, p_4(t))$ be a vector that describes the distribution of the population of bacteria, with $p_i(t)$ the population at time t of bacteria that is between $i-1$ and i hours old. We model the change in population over time by $p(t+1) = Ap(t)$ where A is given by:

$$A = \begin{pmatrix} b_1 & b_2 & b_3 & b_4 \\ s_1 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 \\ 0 & 0 & s_3 & 0 \end{pmatrix}$$

Here $b_i > 0$ and $0 < s_i < 1$ represent the birthrate and the survival rates during the $i^{\rm th}$ time period. In the above model, we assumed that no bacteria lives more than four hours. Notice that the A matrix is non-negative. Model the birthrates and survival rates by:

$$b_i = \gamma_i c_1^{\alpha_1} c_2^{\alpha_2}, \ \ s_i = \delta_i c_1^{\beta_1} c_2^{\beta_2}, \ \ i = 1, \ldots, 4$$

where c_1, c_2 represent some environmental conditions that can be controlled (eg. concentration of a chemical), α_1, α_2, β_1, β_2, γ_i, δ_i are given constants. This model is frequently used in chemistry due to its simplicity. The constants can be found, for example, by taking logarithms in the above equations and then using least-squares estimation.

Our goal is to devise the concentrations c_1, c_2 so that the population of bacteria is depleted as fast as possible, while having constraints on the amounts in c_1, c_2. The decay rate of the population is proportional to the rate of decay of A^k, where A is a non-negative matrix. Hence, we can use the result on Perron-Frobenius eigenvalue, and formulate an optimization problem as follows:

$$\begin{array}{ll} \min_{\lambda,s,v,c} & \lambda \\ \text{s.t.} & b_1 v_1 + b_2 v_2 + b_3 v_3 + b_4 v_4 \leq \lambda v_1, \\ & s_1 v_1 \leq \lambda v_1 \\ & s_2 v_2 \leq \lambda v_3 \\ & s_3 v_3 \leq \lambda v_4 \\ & \frac{1}{2} \leq c_i \leq 2, \ \ b_i = \gamma_i c_1^{\alpha_1} c_2^{\alpha_2}, \ \ s_i = \delta_i c_1^{\beta_1} c_2^{\beta_2}, \ \ i = 1, 2 \end{array}$$

With appropriate scaling of the above equations and inequalities, the problem can be formulated as a GP. The Perron-Frobenius theory is also used in Markov chains, where the matrix representing transition probabilities to the states is non-negative.

Power control in communications. We have n transmitters, labeled $1, \ldots, n$, which transmit at (positive) power levels P_i, $i = 1, \ldots, n$, which are the variables in our power control problem.We also have n receivers, labeled $1, \ldots, n$; receiver i is meant to receive the signal from transmitter i. The power received from transmitter j, at receiver i, is given by $G_{ij}P_j$. Here G_{ij}, which is positive, represents the path gain from transmitter j to receiver i. The signal power at receiver i is $G_{ii}P_i$, and the interference power at receiver i is $\sum_{k\neq i} G_{ik}P_k$.

The noise power at receiver i is given by σ_i. The signal to interference and noise ratio (SINR) of the i-th receiver/transmitter pair is given by

$$S_i = \frac{G_{ii}P_i}{\sigma_i + \sum_{k\neq i} G_{ik}P_k}.$$

We require that the SINR of any receiver/transmitter pair is at least a given threshold $S_{\min}$, and we impose upper and lower bounds onthe transmitter powers. The problem of minimizing the total transmitter power, subject to these constraints, can be expressed as

$$\min_P \sum_{i=1}^n P_i \ : \ P_{\min} \leq P \leq P_{\max}, \ \ \frac{\sigma_i + \sum_{k\neq i} G_{ik}P_k}{G_{ii}P_i} \leq 1/S_{\min}, \ \ i = 1, \ldots, n.$$

This allows to solve the power control problem via GP.

Of course, the above problem can be solved as an LP, but the GP formulation allows us to handle the more general case in which the receiver interference power is any posynomial function of the powers. For example, interference contributions from intermodulation products created by nonlinearities in the receivers typically scale as polynomials in the powers. Indeed, third order intermodulation power from the first and second transmitted signals scales as $P_1P_2^2$ or $P_1^2P_2$; if terms like these are added to the interference power, the power allocation problem above is still a GP.

As an example of a non-trivial extension, where the LP approach fails but the GP model still works, is as follows[3]. Ignore the receiver noise, and assume that the power received from transmitter j to receiver i is $G_{ij}F_jP_j$, where F_j is a random variable that is exponentially distributed, with unit mean. We would like now to ensure that the *outage probabilities*

$$O_i := \mathbf{Prob}\left(G_{ii}F_iP_i \leq \alpha \sum_{j\neq i} G_{ij}F_jP_j\right)$$

are low, where $\alpha > 0$ is a given threshold. (This ensures a "quality of service", or QoS, constraint.) Small values for O_i's mean that the signal powers have a low probability to be small with respect to the interference powers. It turns out that, under the assumption that the F_i's are independent and exponentially distibuted, the outage probabilities have the following closed-form expression[4]:

$$O_i = 1 - \Pi_{j\neq i} \frac{1}{1 + \alpha g_{ij}P_j/P_i}, \ \ g_i := G_{ij}/G_{ii}.$$

[3]See `http://www.stanford.edu/~boyd/papers/outage.html`.

[4]The proof is elementary, look at $\mathbf{Prob}(z_1 > \sum_{i=2}^n z_j)$ when z_j's are independent and exponentially distributed.

Using this, we can minimize the total power subject that each transmitter/receiver pair has an outage probability level less than ϵ via the problem

$$\min_P \; P_1 + \ldots + P_n \;:\; P > 0, \;\; (1-\epsilon)\Pi_{j\neq i}(1 + \alpha g_{ij} P_j/P_i) \leq 1, \;\; i = 1, \ldots, n,$$

which is a GP.

3.10.3 Generalized geometric programming

Several extensions exist for GPs, and the ones below can be cast as GPs (with additional variables and constraints).

Fractional powers of posynomials. Let f_1, f_2 be two posynomials, and let $\alpha_i \geq 0$, $i = 1, 2$. A constraint such as

$$f_1(x)^{\alpha_1} + f_2(x)^{\alpha_2} \leq 1$$

is not a posynomial (unless α_i's are integers). However, we can represent it in GP-compatible format, as

$$t_1^{\alpha_1} + t_2^{\alpha_2} \leq 1, \;\; t_i^{-1} f_i(x) \leq 1, \;\; i = 1, 2.$$

Composition. Let g be a non-decreasing posynomial (all exponents of g are ≥ 0). If $f_1, \ldots, f_K$'s are posynomials, then $g \circ f$ is not a posynomial in general. The constraint

$$g(f_1(x), \ldots, f_K(x)) \leq 1$$

can be represented as

$$g(t_1, \ldots, t_K) \leq 1, \;\; t_i^{-1} f_i(x) \leq 1, \;\; i = 1, \ldots, K.$$

Maximum of posynomials. Likewise, we can cope with $\max(f_1, f_2)$ where f_i's are posynomials, using the representation

$$t_i \leq 1, \;\; t_i^{-1} f_i(x) \leq 1, \;\; i = 1, 2.$$

Generalized posynomials. Generalized posynomials are functions obtained from posynomials by composition with non-decreasing posynomials, addition, maximum and multiplication. As we have seen, such functions can be handled by GPs.

Chapter 4

Duality

4.1 Weak Duality

4.1.1 Primal problem

In this section, we consider a possibly *non-convex* optimization problem

$$p^* := \min_x \ f_0(x) \ : \quad f_i(x) \leq 0, \quad i = 1, \cdots, m, \tag{4.1}$$

where the functions $f_0, f_1, \ldots, f_m$ We denote by $\mathcal{D}$ the domain of the problem (which is the intersection of the domains of all the functions involved), and by $\mathcal{X} \subseteq \mathcal{D}$ its feasible set.

We will refer to the above as the *primal* problem, and to the decision variable x in that problem, as the *primal variable*. One purpose of Lagrange duality is to find a lower bound on a minimization problem (or an upper bounds for a maximization problem). Later, we will use duality tools to derive optimality conditions for convex problems.

4.1.2 Dual problem

Lagrangian. To the problem we associate the Lagrangian $\mathcal{L} : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}$, with values

$$\mathcal{L}(x, \lambda) := f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x)$$

The variables $\lambda \in \mathbf{R}^m$, are called *Lagrange multipliers*.

We observe that, for every feasible $x \in \mathcal{X}$, and every $\lambda \geq 0$, $f_0(x)$ is bounded below by $\mathcal{L}(x, \lambda)$:

$$\forall\, x \in \mathcal{X}, \ \ \forall\, \lambda \in \mathbf{R}^m_+ \ : \ f_0(x) \geq \mathcal{L}(x, \lambda). \tag{4.2}$$

The Lagrangian can be used to express the primal problem (4.1) as an *unconstrained* one. Precisely:

$$p^* = \min_x \ \max_{\lambda \geq 0,\, \nu} \ \mathcal{L}(x, \lambda), \tag{4.3}$$

where we have used the fact that, for any vectors $f \in \mathbf{R}^m$, we have

$$\max_{\lambda \geq 0} \ \lambda^T f = \begin{cases} 0 & \text{if } f \leq 0 \\ +\infty & \text{otherwise.} \end{cases}$$

Lagrange dual function. We then define the *Lagrange dual function* (dual function for short) the function

$$g(\lambda) := \min_x \ \mathcal{L}(x, \lambda).$$

Note that, since g is the pointwise minimum of affine functions ($\mathcal{L}(x, \cdot)$ is affine for every x), it is concave. Note also that it may take the value $-\infty$.

From the bound (4.2), by minimizing over x in the right-hand side, we obtain

$$\forall\, x \in \mathcal{X},\ \ \forall\, \lambda \geq 0\ :\ f_0(x) \geq \min_{x'}\ \mathcal{L}(x', \lambda,) = g(\lambda),$$

which, after minimizing over x the left-hand side, leads to the lower bound

$$\forall\, \lambda \in \mathbf{R}^m_+,\ \ \nu \in \mathbf{R}^p\ :\ p^* \geq g(\lambda).$$

Lagrange dual problem. The best lower bound that we can obtain using the above bound is $p^* \geq d^*$, where

$$d^* = \max_{\lambda \geq 0,\, \nu}\ g(\lambda).$$

We refer to the above problem as the dual problem, and to the vector $\lambda \in \mathbf{R}^m$ as the dual variable. The dual problem involves the maximization of a concave function under convex (sign) constraints, so it is a convex problem. The dual problem always contains the implicit constraint $\lambda \in \mathbf{dom}\, g$.

We have obtained:

Theorem 3 (Weak duality) *For the general (possibly non-convex) problem (4.1),weak duality holds: $p^* \geq d^*$.*

Case with equality constraints. If equality constraints are present in the problem, we can represent them as two inequalities. It turns out that this leads to the same dual, as if we would directly use a single dual variable for each equality constraint, which is not restricted in sign. To see this, consider the problem

$$\begin{array}{rll} p^* := \min\limits_x\ f_0(x)\ : & f_i(x) \leq 0, & i = 1, \cdots, m, \\ & h_i(x) = 0, & i = 1, \cdots, p. \end{array}$$

We write the problem as

$$\begin{array}{rll} p^* := \min\limits_x\ f_0(x)\ : & f_i(x) \leq 0, & i = 1, \cdots, m, \\ & h_i(x) \leq 0,\ -h_i(x) \leq 0, & i = 1, \cdots, p. \end{array}$$

Using a mulitplier $\nu_i^{\pm}$ for the constraint $\pm h_i(x) \leq 0$, we write the associated Lagrangian as

$$\begin{aligned} \mathcal{L}(x, \lambda, \nu^+, \nu^-) &= f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i^+ h_i(x) + \sum_{i=1}^p \nu_i^- (-h_i(x)) \\ &= f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x), \end{aligned}$$

where $\nu := \nu^+ - \nu^-$ does not have any sign constraints.

Thus, inequality constraints in the original problem are associated with sign constraints on the corresponding multipliers, while the multipliers for the equality constraints are not explicitly constrained.

4.1.3 Geometric interpretation

Assume that there is only one inequality constraint in (4.1) ($m = 1$), and let

$$\mathcal{G} := \{(f_1(x), f_0(x)) \ : \ x \in \mathbf{R}^n\}\,.$$

We have

$$p^* = \min_{u,t} \ t \ : \ (u,t) \in \mathcal{G}, \ \ u \leq 0.$$

and

$$g(\lambda) = \min_{u,t} \ (\lambda, 1)^T (u,t) \ : \ (u,t) \in \mathcal{G}.$$

If the minimum is finite, then the inequality $(\lambda, 1)^T (u,t) \geq g(\lambda)$ defines a supporting hyperplane, with slope $-\lambda$, of $\mathcal{G}$ at (u,t). (See Figs. 5.3 and 5.4 in [BV,p.233].)

4.1.4 Minimax inequality

Weak duality can also be obtained as a consequence of the following *minimax inequality*, which is valid for *any* function ϕ of two vector variables x, y, and any subsets $\mathcal{X}$, $\mathcal{Y}$:

$$\max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \ \phi(x,y) \leq \min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \ \phi(x,y). \tag{4.4}$$

To prove this, start from

$$\forall\, x, y \ : \ \min_{x' \in \mathcal{X}} \ \phi(x', y) \leq \max_{y' \in \mathcal{Y}} \ \phi(x, y').$$

and take the minimum over $x \in \mathcal{X}$ on the right-hand side, then the maximum over $y \in \mathcal{Y}$ on the left-hand side.

Weak duality is indeed a direct consequence of the above. To see this, start from the unconstrained formulation (4.3), and apply the above inequality with $\phi = \mathcal{L}$ the Lagrangian of the original problem, and $y = \lambda$ the vector of Lagrange multipliers.

Interpretation as a game. We can interpret the minimax inequality result in the context of a one-shot, zero-sum game. Assume that you have two players A and B, where A controls the decision variable x, while B controls y. We assume that both players have full knowledge of the other player's decision, once it is made. The player A seeks to minimize a payoff (to player B) $\mathcal{L}(x,y)$, while B seeks to maximize that payoff. The right-hand side in (4.4) is the optimal pay-off if the first player is required to play first. Obviously, the first player can do better by playing second, since then he or she knows the opponent's choice and can adapt to it.

4.1.5 Examples

Linear optimization. Consider the LP in standard inequality form

$$p^* = \min_x \ c^T x \ : \ Ax \leq b,$$

where $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$, and the inequality in the constraint $Ax \leq b$ is interpreted component-wise.

The Lagrange function is

$$\mathcal{L}(x, \lambda) = c^T x + \lambda^T (Ax - b)$$

and the corresponding dual function is

$$g(\lambda) = \min_x \mathcal{L}(x,\lambda) = \begin{cases} -b^T\lambda & \text{if } A^T\lambda + c = 0 \\ +\infty & \text{otherwise.} \end{cases}$$

The dual problem reads

$$d^* = \max_\lambda g(\lambda) = \max_\lambda -b^T\lambda \ : \ \lambda \geq 0, \ \ A^T\lambda + c = 0.$$

The dual problem is an LP in standard (sign-constrained) form, just as the primal problem was an LP in standard (inequality) form.

Weak duality implies that

$$c^T x + b^T\lambda \geq 0$$

for every x, λ such that $Ax \leq b$, $A^T\lambda = -c$. This property can be proven directly, by replacing c by $-A^T\lambda$ in the left-hand side of the above inequality, and exploiting $Ax \leq b$ and $\lambda \geq 0$.

We can also consider an LP in standard form:

$$p^* = \min_x c^T x \ : \ Ax = b, \ \ x \geq 0.$$

The equality constraints are associated with a dual variable ν that is not constrained in the dual problem.

The Lagrange function is

$$\mathcal{L}(x,\lambda,\nu) = c^T x - \lambda^T x + \nu^T(b - Ax)$$

and the corresponding dual function is

$$g(\lambda) = \min_x \mathcal{L}(x,\lambda,\nu) = \begin{cases} b^T\nu & \text{if } c = A^T\nu + \lambda \\ +\infty & \text{otherwise.} \end{cases}$$

The dual problem reads

$$d^* = \max_{\lambda \geq 0,\, \nu} g(\lambda,\nu) = \max_\nu b^T\nu \ : \ c \geq A^T\nu.$$

This is an LP in inequality form.

Minimum Euclidean distance problem Consider the problem of minimizing the Euclidean distance to a given affine space:

$$\min \ \frac{1}{2}\|x\|_2^2 \ : \ Ax = b, \tag{4.5}$$

where $A \in \mathbf{R}^{p\times n}$, $b \in \mathbf{R}^p$. We assume that A is full row rank, or equivalently, $AA^T \succ 0$. The Lagrangian is

$$\mathcal{L}(x,\nu) = \frac{1}{2}\|x\|_2^2 + \nu^T(Ax - b),$$

and the Lagrange dual function is

$$g(\nu) = \min_x \mathcal{L}(x,\nu) = \min_x \ \frac{1}{2}\|x\|_2^2 + \nu^T(Ax - b).$$

In this example, the dual function can be computed analytically, using the optimality condition $\nabla_x\mathcal{L}(x,\nu) = x + A^T\nu = 0$. We obtain $x = -A^T\nu$, and

$$g(\nu) = -\frac{1}{2}\nu^T AA^T\nu - b^T\nu.$$

The dual problem expresses as

$$d^* = \max_{\nu} \; g(\nu) = \max_{\nu} \; -\frac{1}{2}\nu^T A A^T \nu - b^T \nu.$$

The dual problem can also be solved analytically, since it is unconstrained (the domain of g is the entire space $\mathbf{R}^p$). We obtain $\nu^* = -(AA^T)^{-1}b$, and

$$d^* = \frac{1}{2} b^T (AA^T)^{-1} b.$$

We have thus obtained the bound $p^* \geq d^*$.

A non-convex boolean problem For a given matrix $W = W^T \succ 0$, we consider the problem

$$p^* = \max_{x} \; x^T W x \;:\; x_i^2 \leq 1, i = 1, \ldots, n.$$

In this maximization problem, Lagrange duality will provide an upper bound on the problem. This is called a "relaxation", as we go above the true maximum, as if we'd relax (ignore) constraints.

The Lagrangian writes

$$\mathcal{L}(x, \lambda) = x^T W x + \sum_{i=1}^{n} \lambda_i (1 - x_i^2) = \mathbf{Tr}\, D_\lambda + x^T (W - D_\lambda) x.$$

where $D_\lambda := \mathbf{diag}(\lambda)$.

To find the dual function, we need to maximize the Lagrangian with respect to the primal variable x. We express this problem as

$$g(\lambda) = \max_{x} \; \mathcal{L}(x, \lambda) = \min_{t} \; t \;:\; \forall\, x, \;\; t \geq \mathbf{Tr}\, D_\lambda + x^T (W - D_\lambda) x.$$

The last inequality holds if and only if

$$\begin{pmatrix} D_\lambda - W & 0 \\ 0 & t - \mathbf{Tr}\, D_\lambda \end{pmatrix} \succeq 0.$$

Hence the dual function is the optimal value of an SDP in one variable:

$$g(\lambda) = \min_{t} \; t \;:\; \begin{pmatrix} D_\lambda - W & 0 \\ 0 & t - \mathbf{Tr}\, D_\lambda \end{pmatrix} \succeq 0.$$

We can solve this problem explicitly:

$$g(\lambda) = \begin{cases} \mathbf{Tr}\, D_\lambda & \text{if } D_\lambda \succeq W \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem involves minimizing (that is, getting the best upper bound) the dual function over the variable $\lambda \geq 0$:

$$d^* = \min_{\lambda} \; \lambda^T \mathbf{1} \;:\; \mathbf{diag}(\lambda) \succeq W.$$

The above is an SDP, in variable λ. Note that $\lambda > 0$ is automatically enforced by the PSD constraint.

Geometric interpretation: The Lagrange relaxation of the primal problem can be interpreted geometrically, as follows. For $t > 0$, $\lambda > 0$, consider the ellipsoids

$$\mathcal{E}_t = \left\{ x \;:\; x^T W x \leq t \right\}, \;\; \mathcal{E}_\lambda = \left\{ x \;:\; x^T D_\lambda x \leq \mathbf{Tr}\, D_\lambda \right\}.$$

The primal problem amounts to find the smallest $t \geq 0$ for which the ellipsoid $\mathcal{E}_t$ contains the ball $\mathcal{B}_\infty := \{x \;:\; \|x\|_\infty \leq 1\}$. Note that for every $\lambda > 0$, $\mathcal{E}_\lambda$ contains the ball $\mathcal{B}_\infty$. To find an upper bound on the problem, we can find the smallest t for which there exist $\lambda > 0$ such that $\mathcal{E}_t \supseteq \mathcal{E}_\lambda$. The latter condition is precisely $\mathbf{diag}(\lambda) \succeq W$, $t \geq \mathbf{Tr}\, D_\lambda$.

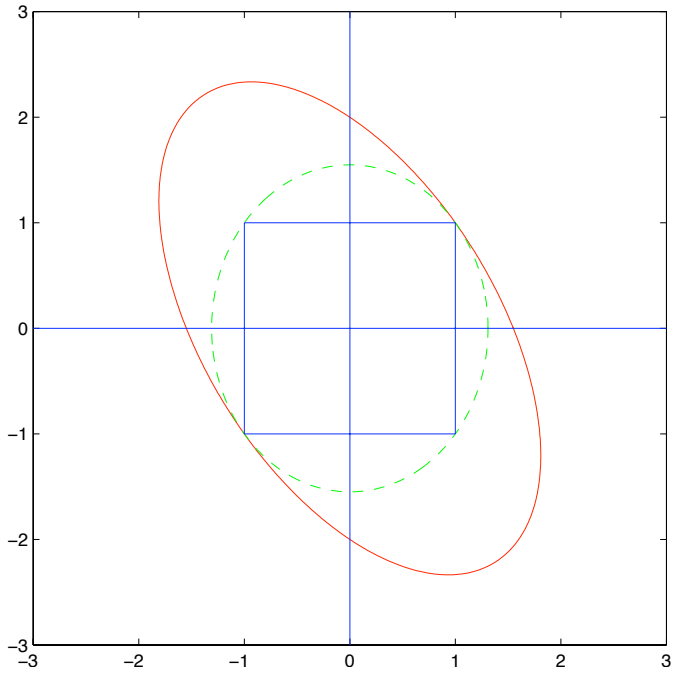

Figure 4.1: Geometric interpretation of dual problem in the boolean quadratic problem.

4.1.6 Semidefinite optimization problem

Consider the SDP in standard form:

$$\max_X \langle C, X\rangle \ : \ \langle A_i, X\rangle = b_i, \ \ i = 1, \ldots, m, \ \ \lambda_{\min}(-X) \leq 0, \tag{4.6}$$

where we have used the minimum eigenvalue function

$$\lambda_{\min}(Z) := \min_Y \langle Y, Z\rangle \ : \ Y \succeq 0, \ \ \mathbf{Tr}\, Y = 1 \tag{4.7}$$

to represent the positive semi-definiteness condition $X \succeq 0$ in the SDP. The proof of this result can be obtained by first showing that we can without loss of generality assume that Z is diagonal, and noticing that we can then restrict Y to be diagonal as well. Note that the above representation proves that $\lambda_{\min}$ is concave, so problem (4.11) is convex as written.

The Lagrangian for the problem above is

$$\mathcal{L}(X, \lambda, \nu) = \langle C, X\rangle + \sum_{i=1}^{m} \nu_i (b_i - \langle A_i, X\rangle) + \lambda \cdot \lambda_{\min}(-X).$$

The dual function for this maximization problem is

$$g(\lambda, \nu) := \max_X \mathcal{L}(X, \lambda, \nu).$$

Consider the following subproblem, in which $R = R^T$ and $\lambda \geq 0$ are given:

$$\begin{aligned} G(R,\lambda) &:= \max_X \langle R, X\rangle + \lambda\lambda_{\min}(-X) \\ &= \max_X \min_{Y\succeq 0,\, \mathbf{Tr}\, Y=1} \langle R - \lambda Y, X\rangle \text{ [eqn. (4.12)]} \\ &= \max_X \min_{Y\succeq 0,\, \mathbf{Tr}\, Y=\lambda} \langle R - Y, X\rangle \text{ [replace } Y \text{ with } \lambda Y] \\ &\geq \underline{G}(R,\lambda) := \min_{Y\succeq 0,\, \mathbf{Tr}\, Y=\lambda} \min_X \langle R - Y, X\rangle \text{ [weak duality].} \end{aligned}$$

The lower bound on G writes

$$\underline{G}(R,\lambda) = \begin{cases} 0 & \text{if } \mathbf{Tr}\, R = \lambda \geq 0, \;\; R \succeq 0, \\ +\infty & \text{otherwise.} \end{cases}$$

This shows that $G(R,\lambda)$ itself is $+\infty$ if (R,λ) is not in the domain of $\underline{G}$. Conversely, if $R \succeq 0$, $\mathbf{Tr}\, R = \lambda$, then the lower bound $\underline{G}(R,\lambda) = 0$ is attained by choosing $X = I$, the identity matrix. Thus, $G = \underline{G}$.

Coming back to the Lagrangian, we need to apply our result to $R = C - \sum_{i=1}^m \nu_i A_i$. The dual function is

$$g(\lambda,\nu) = \begin{cases} 0 & \text{if } R := C - \sum_{i=1}^m \nu_i A_i \succeq 0, \;\; \mathbf{Tr}\, R = \lambda \geq 0, \\ +\infty & \text{otherwise.} \end{cases}$$

We obtain the dual problem

$$d^* = \max_{\lambda,\nu,R} \nu^T b \;:\; R = C - \sum_{i=1}^m \nu_i A_i \succeq 0, \;\; \mathbf{Tr}\, R = \lambda \geq 0,$$

or, after elimination of λ, and noticing that $R \succeq 0$ implies $\mathbf{Tr}\, R \geq 0$:

$$d^* = \max_{\nu} \nu^T b \;:\; C \succeq \sum_{i=1}^m \nu_i A_i \tag{4.8}$$

The dual problem is also an SDP, in standard inequality form.

4.2 Strong duality

4.2.1 Definitions

Duality gap. We have seen how Lagrange duality allows to form a convex optimization problem that provides a lower bound on the original (primal) problem. The *duality gap* is the non-negative number $p^* - d^*$.

We say that *strong duality holds* for the problem if the duality gap is zero: $p^* = d^*$.

4.2.2 A strong duality theorem

Slater's condition. We say that the problem satisfies *Slater's condition* if it is *strictly feasible,* that is:

$$\exists\, x_0 \in \mathcal{D} \;:\; f_i(x_0) < 0, \;\; i = 1,\ldots,m, \;\; h_i(x_0) = 0, \;\; i = 1,\ldots,p.$$

We can replace the above by a *weak form* of Slater's condition, where strict feasibility is not required whenever the function f_i is affine.

We then have the

Theorem 4 (Strong duality via Slater condition) *If the primal problem is convex, and satisfies the weak Slater's condition, then strong duality holds, that is,* $p^* = d^*$.

In particular, we conclude that the duality gap of a linear optimization problem is zero whenever it is feasible. (If not, the gap is also zero, as in fact $p^* = d^* = +\infty$.)

Note that there are many other similar results that guarantee a zero duality gap. For example:

Theorem 5 (Linear optimization problems) *If the functions* $f_0, \ldots, f_m, h_1, \ldots, h_p$ *are all affine, then the duality gap is always zero, provided one of the primal or dual problems is feasible.*

4.2.3 Examples

Minimum Euclidean distance problem. The minimum distance to an affine set (4.5) is convex, and satisfies Slater's condition (in fact, strong duality always holds for this convex quadratic problem). Hence, we know that $p^* = d^*$. This allows us to compute the optimal value of the problem analytically: $p^* = d^* = \frac{1}{2} b^T (AA^T)^{-1} b$.

We can also find a corresponding optimal point: for every ν, the point $x(\nu) = -A^T \nu$ achieves the minimum in the definition of the dual function $g(\nu)$. Let us set $x^* := x(\nu^*)$, where $\nu^* = -(AA^T)^{-1} b$ denotes the optimal dual variable. The point $x^* = A^T (AA^T)^{-1} b$ is optimal for the primal problem. Indeed, it is feasible, since $Ax^* = A^T A (AA^T)^{-1} b = b$, and its objective value equals to the optimal value $(1/2)\|x^*\|_2^2 = \frac{1}{2} b^T (AA^T)^{-1} b = d^* = p^*$. Hence, x^* is optimal, as claimed.

Linear optimization problem. The strong duality theorem applies (without qualification), which shows that any LP in inequality form

$$p^* = \min_x \ c^T x \ : \ Ax \leq b,$$

can be *equivalently* written in the dual form, as an LP:

$$p^* = d^* = \max_\lambda \ -b^T \lambda \ : \ \lambda \geq 0, \ \ A^T \lambda + c = 0.$$

The above LP is in standard form, with the number of constraints and variables exchanged.

Duality is another way to convert any LP in inequality form into a standard form, and vice-versa. (The other method, seen in lecture 5, is via the introduction of new variables.)

Semidefinite optimization problem. Return to the SDP (4.11) and its dual (4.13). Here, strong duality does not necessarily hold. If, however, Slater's condition holds, that is, if there exist $X \succ 0$ that satisfies the constraints of the primal problem, then the duality gap is zero.

Observe the phenomenon observed in the case of LPs. The problem we started with was in standard form, and the dual is the inequality form.

Support vector machine classification. Return to the example seen in lecture 5, involved a binary classification problem. Given m data points $x_i \in \mathbf{R}^n$, each of which is associated with a label $y_i \in \{-1, 1\}$, the problem is to find a

hyperplane that separates, as much as possible, the two classes. Let us denote $Z = [y_1 x_1, \ldots, y_m x_m] \in \mathbf{R}^{n \times m}$.

We assume that the classes are linearly separable, in the sense that there exists $w \in \mathbf{R}^n$, $w \neq 0$, and $b \in \mathbf{R}$, such that

$$y_i(w^T x_i + b) \geq 0, \quad i = 1, \ldots, m.$$

We search for a robust hyperplane, in the sense that the above hold for any data point within a sphere of radius ρ, around the given data points x_i. The robust separability condition reads

$$y_i(w^T x_i + b) \geq \rho \|w\|_2, \quad i = 1, \ldots, m.$$

By homogeneity, we can always assume that (w, b) are chosen so that $\rho\|w\|_2 = 1$. Thus, the largest radius attainable, called the *margin,* is $\rho^* = 1/\|w^*\|_2$, with w^* a minimizer for

$$\min_{w,b} \|w\|_2 \ : \ y_i(w^T x_i + b) \geq 1, \quad i = 1, \ldots, m,$$

or, more compactly:

$$\min_{w,b} \|w\|_2 \ : \ Z^T w + by \geq \mathbf{1}.$$

The above is a QP (after squaring the objective), so we know that strong duality holds.

The Lagrange function is

$$\mathcal{L}(w, b, \lambda) = (1/2)\|w\|_2^2 + \lambda^T(\mathbf{1} - Z^T w - by).$$

(We squared the objective without harm, for simplicity of the next derivation.) The dual function can be explicitly computed as follows. Taking the derivative with respect to w yields $w = Z\lambda$, while zeroing out the derivative with respect to b leads to $\lambda^T y = 0$. Hence:

$$g(\lambda) = \min_{w,b} \mathcal{L}(w, b, \lambda) = \begin{cases} \lambda^T \mathbf{1} - \frac{1}{2}\|Z\lambda\|_2^2 & \text{if } \lambda^T y = 0, \\ +\infty & \text{otherwise.} \end{cases}$$

Making the implicit constraints $\lambda \in \mathbf{dom}\, g$ explicit leads to the dual problem

$$d^* = \max_{\lambda} \ \lambda^T \mathbf{1} - \frac{1}{2}\|Z\lambda\|_2^2 \ : \ \lambda \geq 0, \quad \lambda^T y = 0.$$

This is also a QP, just like the primal problem.

We can interpret the dual geometrically, after massaging it a bit. Scale the variable as $\lambda = \alpha\mu$, with $\alpha > 0$, and $\mu^T y = 0$, $\mu \geq 0$, and optimize over $\alpha > 0$. The optimization problem over α is:

$$\max_{\alpha > 0} \ \alpha(\mu^T \mathbf{1}) - \frac{\alpha^2}{2}\|Z\mu\|_2^2 = \frac{(\mathbf{1}^T \mu)^2}{2\|Z\mu\|_2^2}.$$

Note that the numerator cannot be zero, as long as the data is separable, and $\mu^T y = 0$, $\mu \geq 0$.

This shows that the dual problem can be expressed as

$$d^* = \max_{\mu} \ \frac{(\mathbf{1}^T \mu)^2}{2\|Z\mu\|_2^2} \ : \ \mu \geq 0, \quad \mu^T y = 0.$$

Assume that the first k ($k < m$) first indices in y are $+1$, and the rest -1. The problem reads

$$d^* = \max_{\mu} \ \frac{(\mathbf{1}^T \mu)^2}{2\|Z\mu\|_2^2} \ : \ \mu \geq 0, \ \ \sum_{i=1}^{k} \mu_i = \sum_{i=k+1}^{m} \mu_i.$$

We now use homogeneity of the objective, as follows. Without loss of generality, we can impose the constraint $\sum_{i \leq k} \mu_i = \sum_{i>k} \mu_i = 1$. Then, $\mathbf{1}^T \mu = 2$, and

$$d^* = \max_{\mu} \ \frac{2}{\|Z\mu\|_2^2} \ : \ \mu \geq 0, \ \sum_{i=1}^{k} \mu_i = \sum_{i=k+1}^{m} \mu_i.$$

The margin $1/d^* = 1/p^* = \rho^*$ is equal to the inverse of the above:

$$\rho^* = 1/d^* = \min_{\mu} (1/2) \| \sum_{i=1}^{k} \mu_i x_i - \sum_{i=k+1}^{m} \mu_i x_i \|_2 \ : \ \sum_{i=1}^{k} \mu_i = \sum_{i=k+1}^{m} \mu_i = 1, \ \ \mu \geq 0.$$

The above problem can be interpreted as the minimum Euclidean distance between the convex hulls of the two classes. The inverse of the dual (or primal) value is the margin ρ^*, the largest spherical perturbation that the data can tolerate before becoming not separable. Strong duality implies that the margin is half the distance between the two convex hulls.

4.3 Strong duality for convex problems

4.3.1 Primal and dual problems

In this section, we consider a *convex* optimization problem

$$p^* := \min_{x} \ f_0(x) \ : \quad \begin{array}{ll} f_i(x) \leq 0, & i = 1, \cdots, m, \\ h_i(x) = 0, & i = 1, \cdots, p, \end{array} \tag{4.9}$$

where the functions $f_0, f_1, \ldots, f_m$ are convex, and $h_1, \ldots, h_p$ are affine. We denote by $\mathcal{D}$ the domain of the problem (which is the intersection of the domains of all the functions involved), and by $\mathcal{X} \subseteq \mathcal{D}$ its feasible set.

To the problem we associate the Lagrangian $\mathcal{L} : \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R}$, with values

$$\mathcal{L}(x, \lambda, \nu) := f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x).$$

The dual function is $g : \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R}$, with values

$$g(\lambda, \nu) := \min_{x} \ \mathcal{L}(x, \lambda, \nu).$$

The associated dual problem is

$$d^* = \max_{\lambda \geq 0, \, \nu} \ g(\lambda, \nu).$$

4.3.2 Strong duality via Slater's condition

Duality gap and strong duality. We have seen how weak duality allows to form a convex optimization problem that provides a lower bound on the original (primal) problem, even when the latter is non-convex. The *duality gap* is the non-negative number $p^* - d^*$.

We say that *strong duality holds* for problem (4.9) if the duality gap is zero: $p^* = d^*$.

Slater's condition. We say that the problem satisfies *Slater's condition* if it is *strictly feasible,* that is:

$$\exists\, x_0 \in \mathcal{D} \;:\; f_i(x_0) < 0,\;\; i = 1, \ldots, m,\;\; h_i(x_0) = 0,\;\; i = 1, \ldots, p.$$

We can replace the above by a *weak form* of Slater's condition, where strict feasibility is not required whenever the function f_i is affine.

We then have the

Theorem 6 (Strong duality via Slater condition) *If the primal problem (4.9) is convex, and satisfies the weak Slater's condition, then strong duality holds, that is,* $p^* = d^*$.

Note that there are many other similar results that guarantee a zero duality gap. For example:

Theorem 7 (Quadratic convex optimization problems) *If* f_0 *is quadratic convex, and the functions* $f_1, \ldots, f_m, h_1, \ldots, h_p$ *are all affine, then the duality gap is always zero, provided one of the primal or dual problems is feasible. In particular, strong duality holds for any feasible linear optimization problem.*

A counterexample. Convexity alone is not enough to guarantee strong duality. Consider for example the convex problem

$$\min_{x,y>0}\; e^{-x} \;:\; x^2/y \leq 0,$$

with variables x and y, and domain $\mathcal{D} = \{(x, y) \mid y > 0\}$. We have $p^* = 1$. The Lagrangian is $L(x, y, \lambda) = e^{-x} + \lambda x^2/y$, and the dual function is

$$g(\lambda) = \inf_{x,y>0}(e^{-x} + \lambda x^2/y) = \begin{cases} 0 & \lambda \geq 0 \\ -\infty & \lambda < 0, \end{cases}$$

so we can write the dual problem as

$$d^* = \max_{\lambda}\; 0 \;:\; \lambda \geq 0$$

with optimal value $d^\star = 0$. The optimal duality gap is $p^\star - d^\star = 1$. In this problem, Slater's condition is not satisfied, since $x = 0$ for any feasible pair (x, y).

4.3.3 Geometric interpretation

Assume that there is only one inequality constraint in (4.9) ($m = 1$), and let

$$\mathcal{A} := \{(u, t) \;:\; \exists\, x \in \mathbf{R}^n,\;\; u \geq f_1(x),\;\; t \geq f_0(x)\}\,.$$

The problem is feasible if and only if $\mathcal{A}$ intersects the left-half plane. Furthermore, we have

$$p^* = \min_{u,t}\; t \;:\; (u, t) \in \mathcal{A},\;\; u \leq 0.$$

and

$$g(\lambda) = \min_{u,t}\; (\lambda, 1)^T(u, t) \;:\; (u, t) \in \mathcal{A}.$$

If the minimum is finite, then the inequality $(\lambda, 1)^T(u, t) \geq g(\lambda)$ defines a supporting hyperplane, with slope $-\lambda$, of $\mathcal{A}$ at (u, t). (See Figs. 5.3 and 5.4 in [BV,p.233].)

If the problem is convex, then $\mathcal{A}$ is also convex. If Slater's condition holds, then the interior of $\mathcal{A}$ intersects the left-half plane, and strong duality holds. (See Fig. 5.6 in [BV,p.236].)

4.4 Examples

4.4.1 Minimum Euclidean distance problem

The minimum distance to an affine set mentioned in lecture 11 is

$$\min \ \frac{1}{2}\|x\|_2^2 \ : \ Ax = b, \tag{4.10}$$

where $A \in \mathbf{R}^{p\times n}$, $b \in \mathbf{R}^p$. The problem is convex, and satisfies Slater's condition (in fact, strong duality always holds for this convex quadratic problem). Hence, we know that $p^* = d^*$. This allows us to compute the optimal value of the problem analytically: $p^* = d^* = \frac{1}{2}b^T(AA^T)^{-1}b$.

We can also find a corresponding optimal point: for every ν, the point $x(\nu) = -A^T\nu$ achieves the minimum in the definition of the dual function $g(\nu)$. Let us set $x^* := x(\nu^*)$, where $\nu^* = -(AA^T)^{-1}b$ denotes the optimal dual variable. The point $x^* = A^T(AA^T)^{-1}b$ is optimal for the primal problem. Indeed, it is feasible, since $Ax^* = A^TA(AA^T)^{-1}b = b$, and its objective value equals to the optimal value $(1/2)\|x^*\|_2^2 = \frac{1}{2}b^T(AA^T)^{-1}b = d^* = p^*$. Hence, x^* is optimal, as claimed.

4.4.2 Linear optimization problem

Consider the LP in inequality form:

$$p^* = \min_x \ c^Tx \ : \ Ax \leq b,$$

where $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$. Assume that the above problem is feasible, so that strong duality holds. Then the problem can be *equivalently* written in the dual form, as an LP:

$$p^* = d^* = \max_\lambda \ -b^T\lambda \ : \ \lambda \geq 0, \ \ A^T\lambda + c = 0.$$

The above LP is in standard form, with the number of constraints and variables exchanged.

Duality is another way to convert any LP in inequality form into a standard form, and vice-versa. (The other method, seen in lecture 5, is via the introduction of new variables.)

4.4.3 Support vector machine classification

Return to the example seen in lecture 5, which involved a binary classification problem. Given m data points $x_i \in \mathbf{R}^n$, each of which is associated with a label $y_i \in \{-1, 1\}$, the problem is to find a hyperplane that separates, as much as possible, the two classes. Let us denote $Z = [y_1x_1, \ldots, y_mx_m] \in \mathbf{R}^{n\times m}$.

Ideally, we would like to minimize the number of errors on the training set $(x_i, y_i)_{i=1}^m$. This is hard as it involves a non-convex function. An upper bound on the number of errors is provided by the so-called *hinge loss* function

$$L(w, b) := \sum_{i=1}^m (1 - y_i(w^Tx_i + b))_+.$$

We'd also like to control robustness of the resulting linear classifier, and at the same time guarantee unicity. It turns out that these objectives can be achieved via the following problem:

$$\min_{w,b} \ C \cdot L(w, b) + \frac{1}{2}\|w\|_2^2.$$

where $C > 0$ is a parameter that controls the trade-off between robustness and performance on the training set (a greater C encourages performance at the expense of robustness).

The above can be written as a QP, by introducing slack variables:

$$\min_{w,b,v} \ \frac{1}{2}\|w\|_2^2 + C\sum_{i=1}^m v_i \ : \ v \geq 0, \ \ y_i(w^T x_i + b) \geq 1 - v_i, \ \ i = 1, \ldots, m,$$

or, more compactly:

$$\min_{w,b,v} \ \frac{1}{2}\|w\|_2^2 + Cv^T\mathbf{1} \ : \ v \geq 0, \ \ v + Z^T w + by \geq \mathbf{1}.$$

The corresponding Lagrangian is

$$\mathcal{L}(w, b, \lambda, \mu) = \frac{1}{2}\|w\|_2^2 + Cv^T\mathbf{1} + \lambda^T(\mathbf{1} - v - Z^T w - by) - \mu^T v,$$

where $\mu \in \mathbf{R}^m$ corresponds to the sign constraints on v.

The dual function is given by

$$g(\lambda, \mu) = \min_{w,b} \ \mathcal{L}(w, b, \lambda, \mu).$$

We can readily solve for w by taking derivatives, which leads to $w(\lambda, \mu) = Z\lambda$. Taking derivatives with respect to v yields the constraint $C\mathbf{1} = \lambda + \mu$, while taking derivatives with respect to b leads to the dual constraint $\lambda^T y = 0$. We obtain

$$g(\lambda, \mu) = \begin{cases} \lambda^T\mathbf{1} - \frac{1}{2}\|Z\lambda\|_2^2 & \text{if } \lambda^T y = 0, \ \ \lambda + \mu = C\mathbf{1}, \\ +\infty & \text{otherwise.} \end{cases}$$

We obtain the dual problem

$$d^* = \max_{\lambda \geq 0, \ \mu \geq 0} \ g(\lambda, \mu) = \max_{\lambda} \ \lambda^T\mathbf{1} - \frac{1}{2}\lambda^T Z^T Z\lambda \ : \ 0 \leq \lambda \leq C\mathbf{1}, \ \ \lambda^T y = 0.$$

Strong duality holds, since the primal problem is a QP.

Note that the result depends only on the so-called *kernel matrix* $K = Z^T Z \in \mathcal{S}_+^m$, and the dual problem involves only m variables and m constraints. Hence, the only dependence on the number of dimensions (features), n, is via the required computation of the kernel matrix, that is, on scalar products $x_i^T x_j$, $1 \leq i \leq j \leq m$. Thus, duality allows a great reduction in the computational effort, compared to solving the original QP in n variables and m constraints. This is known as the "kernel trick".

Note also that duality allows to show that the optimal value of the problem is a convex function of the kernel matrix, which allows to optimize over it. We will elaborate on this later.

4.4.4 Semidefinite optimization problem

Consider the SDP in standard form:

$$\max_X \ \langle C, X\rangle \ : \ \langle A_i, X\rangle = b_i, \ \ i = 1, \ldots, m, \ \ \lambda_{\min}(-X) \leq 0, \tag{4.11}$$

where we have used the minimum eigenvalue function

$$\lambda_{\min}(Z) := \min_Y \ \langle Y, Z\rangle \ : \ Y \succeq 0, \ \ \mathbf{Tr}\, Y = 1 \tag{4.12}$$

to represent the positive semi-definiteness condition $X \succeq 0$ in the SDP[1].The proof of this result can be obtained by first showing that we can without loss of generality assume that Z is diagonal, and noticing that we can then restrict Y to be diagonal as well. Note that the above representation proves that $\lambda_{\min}$ is concave, so problem (4.11) is convex as written.

The Lagrangian for the problem above is

$$\mathcal{L}(X, \lambda, \nu) = \langle C, X \rangle + \sum_{i=1}^{m} \nu_i (b_i - \langle A_i, X \rangle) + \lambda \cdot \lambda_{\min}(-X).$$

The dual function for this maximization problem is

$$g(\lambda, \nu) := \max_X \mathcal{L}(X, \lambda, \nu).$$

Consider the following subproblem, in which $R = R^T$ and $\lambda \geq 0$ are given:

$$\begin{aligned} G(R, \lambda) &:= \max_X \langle R, X \rangle + \lambda \lambda_{\min}(-X) \\ &= \max_X \min_{Y \succeq 0,\, \mathbf{Tr}\, Y = 1} \langle R - \lambda Y, X \rangle \text{ [eqn. (4.12)]} \\ &= \max_X \min_{Y \succeq 0,\, \mathbf{Tr}\, Y = \lambda} \langle R - Y, X \rangle \text{ [replace } Y \text{ with } \lambda Y] \\ &\geq \underline{G}(R, \lambda) := \min_{Y \succeq 0,\, \mathbf{Tr}\, Y = \lambda} \min_X \langle R - Y, X \rangle \text{ [weak duality].} \end{aligned}$$

The lower bound on G writes

$$\underline{G}(R, \lambda) = \begin{cases} 0 & \text{if } \mathbf{Tr}\, R = \lambda \geq 0, \;\; R \succeq 0, \\ +\infty & \text{otherwise.} \end{cases}$$

This shows that $G(R, \lambda)$ itself is $+\infty$ if (R, λ) is not in the domain of $\underline{G}$. Conversely, if $R \succeq 0$, $\mathbf{Tr}\, R = \lambda$, then the lower bound $\underline{G}(R, \lambda) = 0$ is attained by choosing $X = I$, the identity matrix. Thus, $G = \underline{G}$.

Coming back to the Lagrangian, we need to apply our result to $R = C - \sum_{i=1}^{m} \nu_i A_i$. The dual function is

$$g(\lambda, \nu) = \begin{cases} 0 & \text{if } R := C - \sum_{i=1}^{m} \nu_i A_i \succeq 0, \;\; \mathbf{Tr}\, R = \lambda \geq 0, \\ +\infty & \text{otherwise.} \end{cases}$$

We obtain the dual problem

$$d^* = \max_{\lambda, \nu, R} \nu^T b \;:\; R = C - \sum_{i=1}^{m} \nu_i A_i \succeq 0, \;\; \mathbf{Tr}\, R = \lambda \geq 0,$$

or, after elimination of λ, and noticing that $R \succeq 0$ implies $\mathbf{Tr}\, R \geq 0$:

$$d^* = \max_{\nu} \nu^T b \;:\; C \succeq \sum_{i=1}^{m} \nu_i A_i \tag{4.13}$$

The dual problem is also an SDP, in standard inequality form.

4.5 Minimax equality theorems

4.5.1 Minimax inequality

As seen in lecture 11, weak duality can be obtained as a consequence of the *minimax inequality*, valid for *any* function ϕ of two vector variables x, y, and

[1] Recall our definition of the scalar product between two symmetric matrices X, Y: $\langle X, Y \rangle = \mathbf{Tr}\, XY$.

any subsets $\mathcal{X}$, $\mathcal{Y}$:

$$d^* := \max_{y\in\mathcal{Y}} \min_{x\in\mathcal{X}} \phi(x,y) \leq \min_{x\in\mathcal{X}} \max_{y\in\mathcal{Y}} \phi(x,y) := p^*. \tag{4.14}$$

Minimax equality theorems identify cases for which the equality $p^* = d^*$ can be proven.

4.5.2 Saddle points

A point $(x^*, y^*) \in \mathcal{X} \times \mathcal{Y}$ is called a *saddle point* if

$$\forall\, x \in \mathcal{X},\ \ \forall\, y \in \mathcal{Y}\ :\ \phi(x^*, y) \leq \phi(x^*, y^*) \leq \phi(x, y^*).$$

The existence of saddle points is related to the minimax equality, as follows:

Proposition 8 *(x^*, y^*) is a saddle point if and only if the minimax equality holds, and is attained, in the sense that*

$$x^* \in \arg\min_{x\in\mathcal{X}} \max_{y\in\mathcal{Y}} \phi(x,y),\ \ y^* \in \arg\max_{y\in\mathcal{Y}} \min_{x\in\mathcal{X}} \phi(x,y).$$

4.5.3 A minimax equality theorem

Theorem 9 (Sion's minimax theorem) *Let $X \subseteq \mathbf{R}^n$ be convex and compact, and let $Y \subseteq \mathbf{R}^m$ be convex. Let $\phi : X \times Y \to \mathbf{R}$ be a function such that for every $y \in Y$, $\phi(\cdot, y)$ is convex and continuous over X, and for every $x \in X$, $\phi(x, \cdot)$ is concave and continuous over Y. Then:*

$$\sup_{y\in Y} \min_{x\in X} \phi(x,y) = \min_{x\in X} \sup_{y\in Y} \phi(x,y).$$

4.5.4 Examples

Support vector machine classification. Return to the example seen in lecture 5, which involved a binary classification problem. Given m data points $x_i \in \mathbf{R}^n$, each of which is associated with a label $y_i \in \{-1, 1\}$, the problem is to find a hyperplane that separates, as much as possible, the two classes. Let us denote $Z = [y_1 x_1, \ldots, y_m x_m] \in \mathbf{R}^{n\times m}$.

We assume that the classes are linearly separable, in the sense that there exists $w \in \mathbf{R}^n$, $w \neq 0$, and $b \in \mathbf{R}$, such that

$$y_i(w^T x_i + b) \geq 0,\ \ i = 1, \ldots, m.$$

We search for a robust hyperplane, in the sense that the above hold for any data point within a sphere of radius ρ, around the given data points x_i. The robust separability condition reads

$$y_i(w^T x_i + b) \geq \rho\|w\|_2,\ \ i = 1, \ldots, m.$$

By homogeneity, we can always assume that (w, b) are chosen so that $\rho\|w\|_2 = 1$. Thus, the largest radius attainable, called the *margin,* is $\rho^* = 1/\|w^*\|_2$, with w^* a minimizer for

$$\min_{w,b} \|w\|_2\ :\ y_i(w^T x_i + b) \geq 1,\ \ i = 1, \ldots, m,$$

or, more compactly:

$$\min_{w,b} \|w\|_2\ :\ Z^T w + by \geq \mathbf{1}.$$

The above is a QP (after squaring the objective), so we know that strong duality holds.

The Lagrange function is

$$\mathcal{L}(w, b, \lambda) = (1/2)\|w\|_2^2 + \lambda^T(\mathbf{1} - Z^T w - by).$$

(We squared the objective without harm, for simplicity of the next derivation.) The dual function can be explicitly computed as follows. Taking the derivative with respect to w yields $w = Z\lambda$, while zeroing out the derivative with respect to b leads to $\lambda^T y = 0$. Hence:

$$g(\lambda) = \min_{w,b} \mathcal{L}(w, b, \lambda) = \begin{cases} \lambda^T \mathbf{1} - \frac{1}{2}\|Z\lambda\|_2^2 & \text{if } \lambda^T y = 0, \\ +\infty & \text{otherwise.} \end{cases}$$

Making the implicit constraints $\lambda \in \mathbf{dom}\, g$ explicit leads to the dual problem

$$d^* = \max_{\lambda} \ \lambda^T \mathbf{1} - \frac{1}{2}\|Z\lambda\|_2^2 \ : \ \lambda \geq 0, \ \ \lambda^T y = 0.$$

This is also a QP, just like the primal problem.

We can interpret the dual geometrically, after massaging it a bit. Scale the variable as $\lambda = \alpha\mu$, with $\alpha > 0$, and $\mu^T y = 0$, $\mu \geq 0$, and optimize over $\alpha > 0$. The optimization problem over α is:

$$\max_{\alpha > 0} \ \alpha(\mu^T \mathbf{1}) - \frac{\alpha^2}{2}\|Z\mu\|_2^2 = \frac{(\mathbf{1}^T\mu)^2}{2\|Z\mu\|_2^2}.$$

Note that the numerator cannot be zero, as long as the data is separable, and $\mu^T y = 0$, $\mu \geq 0$.

This shows that the dual problem can be expressed as

$$d^* = \max_{\mu} \ \frac{(\mathbf{1}^T\mu)^2}{2\|Z\mu\|_2^2} \ : \ \mu \geq 0, \ \ \mu^T y = 0.$$

Assume that the first k ($k < m$) first indices in y are $+1$, and the rest -1. The problem reads

$$d^* = \max_{\mu} \ \frac{(\mathbf{1}^T\mu)^2}{2\|Z\mu\|_2^2} \ : \ \mu \geq 0, \ \ \sum_{i=1}^{k} \mu_i = \sum_{i=k+1}^{m} \mu_i.$$

We now use homogeneity of the objective, as follows. Without loss of generality, we can impose the constraint $\sum_{i \leq k} \mu_i = \sum_{i > k} \mu_i = 1$. Then, $\mathbf{1}^T\mu = 2$, and

$$d^* = \max_{\mu} \ \frac{2}{\|Z\mu\|_2^2} \ : \ \mu \geq 0, \ \ \sum_{i=1}^{k} \mu_i = \sum_{i=k+1}^{m} \mu_i.$$

The margin $1/d^* = 1/p^* = \rho^*$ is equal to the inverse of the above:

$$\rho^* = 1/d^* = \min_{\mu} (1/2)\|\sum_{i=1}^{k} \mu_i x_i - \sum_{i=k+1}^{m} \mu_i x_i\|_2 \ : \ \sum_{i=1}^{k} \mu_i = \sum_{i=k+1}^{m} \mu_i = 1, \ \ \mu \geq 0.$$

The above problem can be interpreted as the minimum Euclidean distance between the convex hulls of the two classes. The inverse of the dual (or primal) value is the margin ρ^*, the largest spherical perturbation that the data can tolerate before becoming not separable. Strong duality implies that the margin is half the distance between the two convex hulls.

4.6 SDP Duality

4.6.1 Primal problem

Consider the SDP in standard form:

$$p^* := \max_X \langle C, X\rangle \ : \ \langle A_i, X\rangle = b_i, \ \ i = 1, \ldots, m, \ \ X \succeq 0, \tag{4.15}$$

where C, A_i are given symmetric matrices, $\langle A, B\rangle = \mathbf{Tr}\, AB$ denotes the scalar product between two symmetric matrices, and $b \in \mathbf{R}^m$ is given.

4.6.2 Dual problem

At first glance, the problem (4.15) is not amenable to the duality theory developed so far, since the constraint $X \succeq 0$ is not a scalar one.

Minimum eigenvalue representation. We develop a dual based on a representation of the problem via the minimum eigenvalue, as

$$p^* = \max_X \langle C, X\rangle \ : \ \langle A_i, X\rangle = b_i, \ \ i = 1, \ldots, m, \ \ \lambda_{\min}(X) \geq 0, \tag{4.16}$$

where we have used the minimum eigenvalue function of a symmetric matrix A, given by

$$\lambda_{\min}(A) := \min_Y \langle Y, A\rangle \ : \ Y \succeq 0, \ \ \mathbf{Tr}\, Y = 1 \tag{4.17}$$

to represent the positive semi-definiteness condition $X \succeq 0$ in the SDP. The proof of the above representation of the minimum eigenvalue can be obtained by first showing that we can without loss of generality assume that A is diagonal, and noticing that we can then restrict Y to be diagonal as well. Note that the above representation proves that $\lambda_{\min}$ is concave, so problem (4.16) is indeed convex as written.

Lagrangian and dual function. The Lagrangian for the *maximization* problem (4.16) is

$$\begin{aligned} \mathcal{L}(X, \lambda, \nu) &= \langle C, X\rangle + \sum_{i=1}^m \nu_i (b_i - \langle A_i, X\rangle) + \lambda \cdot \lambda_{\min}(X) \\ &= \nu^T b + \langle C - \sum_{i=1}^m \nu_i A_i, X\rangle + \lambda \cdot \lambda_{\min}(X), \end{aligned}$$

where $nu \in \mathbf{R}^m$ and $\lambda \geq 0$ are the dual variables. The corresponding dual function

$$g(\lambda, \nu) := \max_X \ \mathcal{L}(X, \lambda, \nu).$$

involves the following subproblem, in which $Z = C - \sum_{i=1}^m \nu_i A_i$ and $\lambda \geq 0$ are given:

$$G(\lambda, Z) := \max_X \langle Z, X\rangle + \lambda \lambda_{\min}(X). \tag{4.18}$$

For fixed $\lambda \geq 0$, the function $G(\cdot, \lambda)$ is the conjugate of the convex function $-\lambda\lambda_{\min}$.

We have

$$\begin{aligned} G(\lambda, Z) &= \max_X \left(\langle Z, X\rangle + \lambda \min_{Y\succeq 0,\, :\, \mathbf{Tr}\, Y=1} \langle Y, X\rangle \right) \\ &= \max_X \min_{Y\succeq 0,\, \mathbf{Tr}\, Y=1} \langle Z + \lambda Y, X\rangle \text{ [eqn. (4.17)]} \\ &= \max_X \min_{Y\succeq 0,\, \mathbf{Tr}\, Y=\lambda} \langle Z + Y, X\rangle \text{ [replace } Y \text{ by } \lambda Y] \\ &\leq \min_{Y\succeq 0,\, \mathbf{Tr}\, Y=\lambda} \max_X \langle Z + Y, X\rangle \text{ [minimax inequality]} \\ &= \min_{Y\succeq 0,\, \mathbf{Tr}\, Y=\lambda} \begin{cases} 0 & \text{if } Z + Y = 0 \\ +\infty & \text{otherwise} \end{cases} \\ &= \underline{G}(\lambda, Z), \end{aligned}$$

where

$$\underline{G}(\lambda, Z) := \begin{cases} 0 & \text{if } \mathbf{Tr}\, Z = -\lambda \leq 0, \;\; Z \preceq 0, \\ +\infty & \text{otherwise.} \end{cases}$$

We now show that $G(\lambda, Z) = \underline{G}(\lambda, Z)$. To prove this, first note that $G(\lambda, Z) \geq 0$ (since $X = 0$ is feasible). This shows that $G(\lambda, Z)$ itself is 0 if (λ, Z) is in the domain of $\underline{G}$. Conversely, if $\mathbf{Tr}\, Z + \lambda \neq 0$, choosing $X = \epsilon t I$ with $\epsilon = \mathbf{sign}(\mathbf{Tr}\, Z + \lambda)$ and $t \to +\infty$ implies $G(\lambda, Z) = +\infty$. Likewise, if $\mathbf{Tr}\, Z + \lambda = 0$ and $\lambda < 0$, we choose $X = tI$ with $t \to +\infty$, with the same result. Finally, if $\mathbf{Tr}\, Z = -\lambda \leq 0$ but $\lambda_{\max}(Z) > 0$, choose $X = tuu^T$, with u a unit-norm eigenvector corresponding to the largest eigenvalue of Z, and $t \to +\infty$. Here, we have

$$\langle Z, X\rangle + \lambda\lambda_{\min}(X) = t(\lambda_{\max}(Z) + \lambda) \to +\infty,$$

where we have exploited the fact that $\lambda_{\max}(Z) + \lambda > \lambda \geq 0$.

Dual problem. Coming back to the Lagrangian, we need to apply our result to $Z = C - \sum_{i=1}^m \nu_i A_i$. The dual function is

$$g(\lambda, \nu) = \begin{cases} 0 & \text{if } Z := C - \sum_{i=1}^m \nu_i A_i \preceq 0, \;\; \mathbf{Tr}\, Z = -\lambda \geq 0, \\ +\infty & \text{otherwise.} \end{cases}$$

We obtain the dual problem

$$d^* = \min_{\lambda,\nu,Z} \; \nu^T b \;:\; Z = C - \sum_{i=1}^m \nu_i A_i \preceq 0, \;\; \mathbf{Tr}\, Z = -\lambda \geq 0,$$

or, after elimination of λ, and noticing that $Z \preceq 0$ implies $\mathbf{Tr}\, Z \leq 0$:

$$d^* = \min_{\nu} \; \nu^T b \;:\; \sum_{i=1}^m \nu_i A_i \succeq C \tag{4.19}$$

The dual problem is also an SDP, in standard inequality form.

4.6.3 Conic approach

Conic Lagrangian. The same dual can be obtained with the "conic" Lagrangian

$$\mathcal{L}(X, \nu, Y) := \langle C, X\rangle + \sum_{i=1}^m \nu_i (b_i - \langle A_i, X\rangle) + \langle Y, X\rangle,$$

where now we associate a *matrix* dual variable Y to the constraint $X \succeq 0$.

Let us check that the Lagrangian above "works", in the sense that we can represent the constrained maximization problem (4.15) as an unconstrained, maximin problem:

$$p^* = \max_X \min_{Y \succeq 0} \mathcal{L}(X, \nu, Y).$$

We need to check that, for an arbitrary matrix Z, we have

$$\min_{Y \succeq 0} \langle Y, X \rangle = \begin{cases} 0 & \text{if } X \succeq 0 \\ -\infty & \text{otherwise.} \end{cases} \tag{4.20}$$

This is an immediate consequence of the following:

$$\min_{Y \succeq 0} \langle Y, X \rangle = \min_{t \geq 0} \min_{Y \succeq 0,\, \mathbf{Tr}\, Y = t} \langle Y, X \rangle = \min_{t \geq 0} t \lambda_{\min}(X),$$

where we have exploited the representation of the minimum eigenvalue given in (4.17). The geometric interpretation is that the cone of positive-semidefinite matrices has a 90^o angle at the origin.

Dual problem. The minimax inequality then implies

$$p^* \leq d^* := \min_{\nu,\, Y \succeq 0} \max_X \mathcal{L}(X, \nu, Y).$$

The corresponding dual function is

$$g(Y, \nu) = \max_X \mathcal{L}(X, \nu, Y) = \begin{cases} \nu^T b & \text{if } C - \sum_{i=1}^m \nu_i A_i + Y = 0 \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem then writes

$$d^* = \min_{\nu,\, Y \succeq 0} g(Y, \nu) = \min_{\nu,\, Y \succeq 0} \nu^T b \ : \ C - \sum_{i=1}^m \nu_i A_i = -Y \preceq 0.$$

After elimination of the variable Y, we find the same problem as before, namely (4.19).

4.6.4 Weak and strong duality

Weak duality. For the maximization problem (4.16), weak duality states that $p^* \leq d^*$. Note that the fact that weak duality inequality

$$\nu^T b \geq \langle C, X \rangle$$

holds for any primal-dual feasible pair (X, ν), is a direct consequence of (4.20).

Strong duality. From Slater's theorem, strong duality will hold if the primal problem is strictly feasible, that is, if there exist $X \succ 0$ such that $\langle A_i, X \rangle = b_i$, $i = 1, \ldots, m$.

Using the same approach as above, one can show that the dual of problem (4.19) is precisely the primal problem (4.16). Hence, if the dual problem is strictly feasible, then strong duality also holds.Recall that we say that a problem is *attained* if its optimal set is not empty. It turns out that if both problems are strictly feasible, then both problems are attained.

A strong duality theorem. The following theorem summarizes our results.

Theorem 10 (strong duality in SDP) *Consider the SDP*

$$p^* := \max_X \langle C, X\rangle \ : \ \langle A_i, X\rangle = b_i, \ \ i = 1, \ldots, m, \ \ X \succeq 0$$

and its dual

$$d^* = \min_\nu \ \nu^T b \ : \ \sum_{i=1}^m \nu_i A_i \succeq C.$$

The following holds:

- *Duality is symmetric, in the sense that the dual of the dual is the primal.*
- *Weak duality always holds: $p^* \leq d^*$, so that, for any primal-dual feasible pair (X, ν), we have $\nu^T b \geq \langle C, X\rangle$.*
- *If the primal (resp. dual) problem is bounded above (resp. below), and strictly feasible, then $p^* = d^*$ and the dual (resp. primal) is attained.*
- *If both problems are strictly feasible, then $p^* = d^*$ and both problems are attained.*

4.6.5 Examples

An SDP where strong duality fails. Contrarily to linear optimization problems, SDPs can fail to have a zero duality gap, even when they are feasible. Consider the example:

$$p^* = \min_x \ x_2 \ : \ \begin{pmatrix} x_2 + 1 & 0 & 0 \\ 0 & x_1 & x_2 \\ 0 & x_2 & 0 \end{pmatrix} \succeq 0.$$

Any primal feasible x satisfies $x_2 = 0$. Indeed, positive-semidefiniteness of the lower-right 2×2 block in the LMI of the above problem writes, using Schur complements, as $x_1 \leq 0$, $x_2^2 \leq 0$. Hence, we have $p^* = 0$. The dual is

$$d^* = \max_{Y \in \mathcal{S}^3} \ -Y_{11} \ : \ Y \succeq 0, \ \ Y_{22} = 0, \ \ 1 - Y_{11} - 2Y_{23} = 0.$$

Any dual feasible Y satisfies $Y_{23} = 0$ (since $Y_{22} = 0$), thus $Y_{11} = -1 = d^*$.

An eigenvalue problem. For a matrix $A \in \mathcal{S}_+^n$, we consider the SDP

$$p^* = \max_X \langle A, X\rangle \ : \ \mathbf{Tr}\, X = 1, \ \ X \succeq 0. \tag{4.21}$$

The associated Lagrangian, using the conic approach, is

$$\mathcal{L}(X, Y, \nu) = \langle A, X\rangle + \nu(1 - \mathbf{Tr}\, X) + \langle Y, X\rangle,$$

with the matrix dual variable $Y \succeq 0$, while $\nu \in \mathbf{R}$ is free.

The dual function is

$$g(Y, \nu) = \max_X \ \mathcal{L}(X, Y, \nu) = \left\{ \begin{array}{ll} \nu & \text{if } \nu I = Y + A \\ +\infty & \text{otherwise.} \end{array} \right.$$

We obtain the dual problem

$$p^* \leq d^* = \min_{\nu, Y} \ \nu \ : \ Y + A = \lambda I, \ \ Y \succeq 0.$$

Eliminating Y leads to

$$d^* = \min_{\nu} \{\nu \ : \ \nu I \succeq A\} = \lambda_{\max}(A).$$

Both the primal and dual problems are strictly feasible, so $p^* = d^*$, and both values are attained. This proves the representation (4.21) for the largest eigenvalue of A.

SDP relaxations for a non-convex quadratic problem. Previously, we have seen two kinds of relaxation for the non-convex problem

$$p^* := \max_x \ x^T W x \ : \ x_i^2 \leq 1, \ \ i = 1, \ldots, n,$$

where the symmetric matrix $W \in \mathcal{S}^n$ is given.

One relaxation is based on a standard relaxation of the constraints, and leads to

$$p^* \leq d^{\text{lag}} := \min_D \ \mathbf{Tr}\, D \ : \ D \succeq W, \ \ D \text{ diagonal.} \tag{4.22}$$

Another relaxation involved expressing the problem as an SDP with rank constraints on the a matrix $X = xx^T$:

$$d^{\text{rank}} := \max_X \langle W, X \rangle \ : \ X \succeq 0, \ \ X_{ii} = 1, \ \ i = 1, \ldots, m.$$

Let us examine the dual of the first relaxation (4.22). We note that the problem is strictly feasible, so strong duality holds. Using the conic approach, we have

$$\begin{aligned} d^{\text{lag}} &:= \min_D \max_{Y \succeq 0} \mathbf{Tr}\, D + \langle Y, W - D \rangle \\ &= \max_{Y \succeq 0} \min_D \mathbf{Tr}\, D + \langle Y, W - D \rangle \\ &= \max_Y \langle Y, W \rangle \ : \ \succeq 0, \ \ Y_{ii} = 1, \ \ i = 1, \ldots, m \\ &= d^{\text{rank}}. \end{aligned}$$

This shows that both Lagrange and rank relaxations give the same value, and are dual of each other.

In general, for arbitrary non-convex quadratic problems, the rank relaxation can be shown to be always better than the Lagrange relaxation, as the former is the (conic) dual to the latter. If either is strictly feasible, then they have the same optimal value.

4.7 SOCP Duality

Second-order cone optimization is a special case of semi-definite optimization. It is, however, instructive to develop a more direct approach to duality for SOCPs.

4.7.1 Conic approach

We start from the second-order cone problem in inequality form:

$$p^* := \min_x \ c^T x \ : \ \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \ \ i = 1, \ldots, m,$$

where $c \in \mathbf{R}^n$, $A_i \in \mathbf{R}^{n_i \times n}$, $b_i \in \mathbf{R}^{n_i}$, $c_i \in \mathbf{R}^n$, $d_i \in \mathbf{R}$, $i = 1, \ldots, m$.

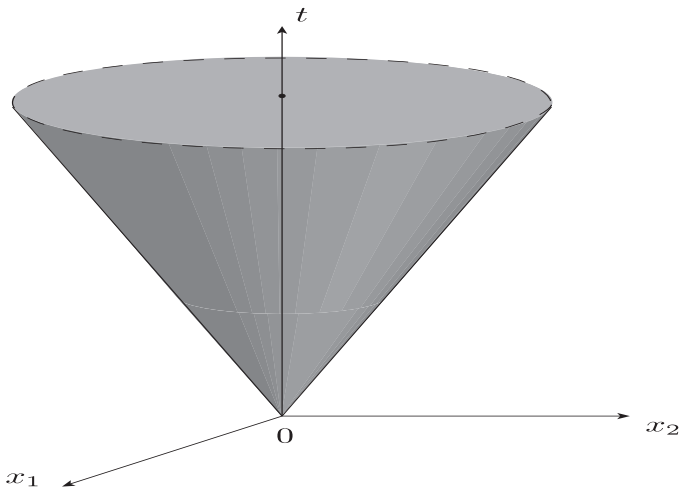

Figure 4.2: Geometric interpretation of the property (4.23). The two orthogonal vectors in black form the maximum angle attainable by vector in the second-order cone. The vector in red forms a greater angle with the vector on the left, and the corresponding scalar product is unbounded.

Conic Lagrangian. To build a Lagrangian for this problem, we use the fact that, for any pair (t, y):

$$\max_{(u,\lambda)\,:\,\|u\|_2\le\lambda} -u^T y - t\lambda = \max_{\lambda\ge 0}\ \lambda(\|y\|_2 - t) = \begin{cases} 0 & \text{if } \|y\|_2 \le t \\ +\infty & \text{otherwise.} \end{cases} \tag{4.23}$$

A geometric interpretation is shown in Fig 4.2. The above means that the second-order cone has a 90^o angle at the origin. To see this, observe that

$$\max_{(u,\lambda)\,:\,\|u\|_2\le\lambda} -u^T y - t\lambda = -\min_{(u,\lambda)\,:\,\|u\|_2\le\lambda} \begin{pmatrix} u \\ \lambda \end{pmatrix}^T \begin{pmatrix} y \\ t \end{pmatrix}.$$

The objective in the right-hand side is proportional to the cosine of the angle between the vectors involved. The largest angle achievable between any two vectors in the second-order cone is 90^o. If $\|y\|_2 > t$, then the cosine reaches negative values, and the maximum scalar product becomes infinite.

Consider the following Lagrangian, with variables x, $\lambda \in \mathbf{R}^m$, $u_i \in \mathbf{R}^{n_i}$, $i = 1, \ldots, m$:

$$\mathcal{L}(x, \lambda, u_1, \ldots, u_m) = c^T x - \sum_{i=1}^{m} \left[u_i^T (A_i x + b_i) + \lambda_i (c_i^T x + d_i)\right].$$

Using the fact above leads to the following minimax representation of the primal problem:

$$p^* = \min_x \max_{\|u_i\|_2\le\lambda_i,\ i=1,\ldots,m} \mathcal{L}(x, \lambda, u_1, \ldots, u_m).$$

Conic dual. Weak duality expresses as $p^* \ge d^*$, where

$$d^* := \max_{\|u_i\|_2\le\lambda_i,\ i=1,\ldots,m} \min_x\ \mathcal{L}(x, \lambda, u_1, \ldots, u_m).$$

The inner problem, which corresponds to the dual function, is very easy to solve as the problem is unconstrained and the objective affine (in x). Setting

the derivative with respect to x leads to the dual constraints

$$c = \sum_{i=1}^{m} [A_i^T u_i + \lambda_i c_i].$$

We obtain

$$d^* = \max_{\lambda, u_i,\ i=1,\ldots,m} -\lambda^T d - \sum_{i=1}^{m} u_i^T b_i \ :\ c = \sum_{i=1}^{m} [A_i^T u_i + \lambda_i c_i],\ \ \|u_i\|_2 \leq \lambda_i,\ \ i = 1, \ldots, m.$$

The above is an SOCP, just like the original one.

Direct approach. As for the SDP case, it turns out that the above "conic" approach is the same as if we had used the Lagrangian

$$\mathcal{L}_{\text{direct}}(x, \lambda) = c^T x + \sum_{i=1}^{m} \lambda_i \left[\|A_i x + b_i\|_2 - (c_i^T x + d_i) \right].$$

Indeed, we observe that

$$\mathcal{L}_{\text{direct}}(x, \lambda) = \max_{u_i,\ i=1,\ldots,m} \mathcal{L}(x, \lambda, u_1, \ldots, u_m) \ :\ \|u_i\|_2 \leq \lambda_i,\ \ i = 1, \ldots, m.$$

4.7.2 Strong duality

Strong duality results are similar to those for SDP: a sufficient condition for strong duality to hold is that one of the primal or dual problems is strictly feasible. If both are, then the optimal value of both problems is attained.

Theorem 11 *Consider the SOCP*

$$p^* := \min_x \ c^T x \ :\ \|A_i x + b_i\|_2 \leq c_i^T x + d_i,\ \ i = 1, \ldots, m,$$

and its dual

$$d^* = \max_{\lambda, u_i,\ i=1,\ldots,m} -\lambda^T d - \sum_{i=1}^{m} u_i^T b_i \ :\ c = \sum_{i=1}^{m} [A_i^T u_i + \lambda_i c_i],\ \ \|u_i\|_2 \leq \lambda_i,\ \ i = 1, \ldots, m.$$

The following holds:

- *Duality is symmetric, in the sense that the dual of the dual is the primal.*
- *Weak duality always holds: $p^* \geq d^*$, so that, for any primal-dual feasible pair $(x, (u_i, \lambda_i)_{i=}^m)$, we have $\lambda^T d + \sum_{i=1}^m u_i^T b_i \leq c^T x$.*
- *If the primal (resp. dual) problem is bounded above (resp. below), and strictly feasible, then $p^* = d^*$ and the dual (resp. primal) is attained.*
- *If both problems are strictly feasible, then $p^* = d^*$ and both problems are attained.*

4.7.3 KKT conditions for SDP

Consider the SDP

$$p^* = \min_x \ c^T x \ :\ F(x) := F_0 + \sum_{i=1}^{m} x_i F_i \succeq 0,$$

with $c \in \mathbf{R}^n$, $F_i \in \mathcal{S}^n$, $i = 0, \ldots, m$. The Lagrangian is

$$\mathcal{L}(x, Y) = c^T x - \mathbf{Tr}\, F(x) Y,$$

and the dual problem reads

$$d^* = \max_{Y \succeq 0} \min_x \mathcal{L}(x, Y) = \max_Y -\mathbf{Tr}\, F_0 Y \ : \ \mathbf{Tr}\, F_i Y = c_i, \ \ i = 1, \ldots, n, \ ; Y \succeq 0.$$

The KKT conditions for optimality are as follows:

1. $F(x) \succeq 0$,
2. $Y \succeq 0$, $\mathbf{Tr}(F_i Y) = c_i$, $i = 1, \ldots, n$,
3. $\mathbf{Tr}(F(x)Y) = 0$.

The last condition can be expressed as $F(x)Y = 0$, according to the following result: Let $F, Y \in \mathcal{S}^n$. If $F \succeq 0$ and $Y \succeq 0$ then $\mathbf{Tr}(FY) = 0$ is equivalent to $FY = 0$.

Proof: Let $Y^{1/2}$ be the square root of Y (the unique positive semi-definite solution to $Z^2 = Y$). We have $\mathbf{Tr}\, FY = \mathbf{Tr}\, \tilde{F} = 0$, where $\tilde{F} := Y^{1/2} F Y^{1/2}$. Since $F \succeq 0$, we have $\tilde{F} \succeq 0$. The trace of $\tilde{F}$ being zero then implies that $\tilde{F} = 0$.

Using the eigenvalue decomposition, we can reduce the problem to the case when Y is diagonal. Let us assume that

$$Y = \begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}, \ \ F = \begin{pmatrix} F_{11} & F_{12} \\ F_{12}^T & F_{22} \end{pmatrix}$$

where $\Lambda \succ 0$ is diagonal, and contains the eigenvalues of Y, and the matrix F_{11} is of the same size as Λ (which is equal to the rank of Y). The condition $\tilde{F} = Y^{1/2} F Y^{1/2} = 0$ expresses as

$$0 = \begin{pmatrix} \Lambda^{1/2} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} F_{11} & F_{12} \\ F_{12}^T & F_{22} \end{pmatrix} \begin{pmatrix} \Lambda^{1/2} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \Lambda^{1/2} F_{11} \Lambda^{1/2} & 0 \\ 0 & 0 \end{pmatrix}.$$

Since $\Lambda \succ 0$, we obtain $F_{11} = 0$. But $F \succeq 0$ then implies $F_{12} = 0$ (use Schur complements), thus

$$FY = \begin{pmatrix} 0 & 0 \\ 0 & F_{22} \end{pmatrix} \begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix} = 0,$$

as claimed. Thus the last KKT condition can be written as $F(x)Y = 0$.

Theorem 12 (KKT conditions for SDP) *The SDP*

$$p^* := \min_x \ c^T x \ : \ F(x) := F_0 + \sum_{i=1}^m x_i F_i \succeq 0$$

admits the dual bound $p^* \geq d^*$, where

$$d^* = \max_Y -\mathbf{Tr}\, F_0 Y \ : \ \mathbf{Tr}\, F_i Y = c_i, \ \ i = 1, \ldots, n, \ \ Y \succeq 0.$$

If both problems are strictly feasible, then the duality gap is zero: $p^* = d^*$, and both values are attained. Then, a pair (x, Y) is primal-dual optimal if and only if the KKT conditions

1. Primal feasibility: $F(x) \succeq 0$,

2. Dual feasibility: $\mathbf{Tr}\, F_i Y = c_i, \ \ i = 1, \ldots, n, \ \ Y \succeq 0$,

3. Complementary slackness: $F(x)Y = 0$,

hold.

Recall LP duality for a problem of the from $\min_x\{c^T x : Ax \leq b\}$ with dual variables y has the KKT conditions were $\forall i : y_i(b - Ax)_i = 0$. This can be compactly written as $\mathbf{diag}(y)\,\mathbf{diag}(b - Ax) = 0$ which is similar to the KKT conditions for SDP (with $Y = \mathbf{diag}(y)$). This should come as no surprise as SDP problems include LP problems as a special case.

4.7.4 Examples

4.7.5 Minimum distance to an affine subspace

Return to the problem seen in lecture 11:

$$p^* = \min \|x\|_2 \ : \ Ax = b, \tag{4.24}$$

where $A \in \mathbf{R}^{p\times n}$, $b \in \mathbf{R}^p$, with b in the range of A. We have seen how to develop a dual when the objective is squared. Here we will work directly with the Euclidean norm.

The above problem is an SOCP. To see this, simply put the problem in epigraph form. Hence the above theory applies. A more direct (equivalent) way, which covers cases when norms appear in the objective, is to use the representation of the objective as a maximum:

$$p^* = \min_x \max_{\nu,\, \|u\|_2\leq 1} x^T u + \nu^T(b - Ax) \geq d^* = \max_{\nu,\, \|u\|_2\leq 1} \min_x x^T u + \nu^T(b - Ax).$$

The dual function is

$$g(u) = \min_x x^T u + \nu^T(b - Ax) = \begin{cases} \nu^T b & \text{if } A^T\nu = u, \\ -\infty & \text{otherwise.} \end{cases}$$

We obtain the dual

$$d^* = \max_{\nu,\, u} b^T\nu \ : \ A^T\nu = u, \ \ \|u\|_2 \leq 1.$$

Eliminating u:

$$d^* = \max_{\nu} b^T\nu \ : \ \|A^T\nu\|_2 \leq 1.$$

4.7.6 Robust least-squares

Consider a least-squares problem

$$\min_x \|Ax - b\|_2,$$

where $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$. In practice, A may be noisy. To handle this, we assume that A is additively perturbed by a matrix bounded in largest singular value norm (denoted $\|\cdot\|$ in the sequel) by a given number $\rho \geq 0$. The robust counterpart to the least-squares problem then reads

$$\min_x \max_{\|\Delta\|\leq\rho} \|(A + \Delta)x - b\|_2.$$

Using convexity of the norm, we have

$$\forall\, \Delta, \ \ \|\Delta\| \leq \rho \ : \ \|(A + \Delta)x - b\|_2 \leq \|Ax - b\|_2 + \|\Delta x\|_2 \leq \|Ax - b\|_2 + \rho\|x\|_2.$$

The upper bound is attained, with the choice[2]

$$\Delta = \frac{\rho}{\|x\|_2 \cdot \|Ax-b\|_2}(Ax-b)x^T.$$

Hence, the robust counterpart is equivalent to the SOCP

$$\min_x \ \|Ax-b\|_2 + \rho\|x\|_2.$$

Again, we can use epigraph representations for each norm in the objective:

$$\min_{x,t,\tau} \ t + \rho\tau \ : \ t \geq \|Ax-b\|_2, \ \ \tau \geq \|x\|_2.$$

and apply the standard theory for SOCP developed in section 4.7.1. Strong duality holds, since the problem is strictly feasible.

An equivalent, more direct approach is to represent each norm as a maximum:

$$p^* = \min_x \max_{\|u\|_2\leq 1, \ \|v\|_2\leq\rho} u^T(b-Ax) + v^Tx.$$

Exchanging the min and the max leads to the dual

$$p^* \geq d^* = \max_{\|u\|_2\leq 1, \ \|v\|_2\leq\rho} \min_x \ u^T(b-Ax) + v^Tx.$$

The dual function is

$$g(u,v) = \min_x \ v^T(b-Ax) + u^Tx = \begin{cases} v^Tb & \text{if } A^Tv+u=0, \\ -\infty & \text{otherwise.} \end{cases}$$

Eliminating u, we obtain the dual

$$d^* = \max_{u,v} \ v^Tb \ : \ \|A^Tv\|_2 \leq 1, \ \ \|v\|_2 \leq \rho.$$

As expected, when ρ grows, the dual solution tends to the least-norm solution to the system $Ax = b$. It turns out that the above approach leads to a dual that is equivalent to the SOCP dual, and that strong duality holds.

4.7.7 Probabilistic classification

Consider a binary classification problem, in which the data points for each class and radom variables x_+, x_-, each assumed to obey to a given Gaussian distribution $\mathcal{N}(\hat{x}_\pm, \Sigma_\pm)$, where $\hat{x}_\pm \in \mathbf{R}^n$, $\Sigma_\pm \in \mathcal{S}^n_{++}$ are the given class-dependent means and covariance matrices, respectively. We seek an hyperplane $\mathcal{H}(w,b) := \{x \ : \ w^Tx + b = 0\}$ that probabilistically separates the two classes, in the sense that

$$\mathbf{Prob}\{x_+ \ : \ (w^Tx_+ + b) \geq 0\} \geq 1-\epsilon, \ \ \mathbf{Prob}\{x_- \ : \ (w^Tx_- + b) \leq 0\} \geq 1-\epsilon,$$

where ϵ is a given small number. The interpretation is that we would like that, with high probability, the samples taken from each distribution fall on the correct side of the hyperplane. The number $\epsilon < 1/2$ allows to set the level of reliability of the classification, with small ϵ corresponding to a low probability of mis-classification. We assume that $\hat{x}_+ \neq \hat{x}_-$.

When x obeys to a distribution $\mathcal{N}(\hat{x}, \Sigma)$, the random variable w^Tx+b follows the distribution $\mathcal{N}(\hat{\xi}, \sigma^2)$, with $\hat{\xi} := w^T\hat{x} + b$, $\sigma^2 = w^T\Sigma w$. We can write

[2] We assume that $x \neq 0$, $Ax \neq b$. These cases are easily analyzed and do not modify the result.

$\xi = \hat{\xi} + \sigma u$, with u a normal (zero-mean, unit variance) random variable. We thus have

$$\mathbf{Prob}\{\xi \,:\, \xi \geq 0\} = \mathbf{Prob}\{u \,:\, u \geq -\hat{\xi}/\sigma\} = 1 - \Phi(-\hat{\xi}/\sigma),$$

where Φ is the cumulative density function of the normal distribution, namely $\Phi(\alpha) := \mathbf{Prob}\{u \,:\, u \leq \alpha\}$. Since Φ is monotone increasing, the condition $\mathbf{Prob}\{\xi \,:\, \xi \geq 0\} \geq 1 - \epsilon$ is equivalent to $-\hat{\xi}/\sigma \leq \Phi^{-1}(\epsilon)$, or $\hat{\xi} \geq \kappa_\epsilon \sigma$, where $\kappa_\epsilon := -\Phi^{-1}(\epsilon)$. Note that $\kappa_\epsilon > 0$ whenever $0 \leq \epsilon < 1/2$.

We obtain that the probability constraints above write

$$w^T \hat{x}_+ + b \geq \kappa_\epsilon \sqrt{w^T \Sigma_+ w}, \;\; w^T \hat{x}_- + b \leq -\kappa_\epsilon \sqrt{w^T \Sigma_- w}.$$

Note that when $\Sigma_\pm = 0$, he above simply requires the correct classification of the means $\hat{x}_\pm$. When $\Sigma_\pm$ grows in size, he conditions become more and more stringent. We can eliminate b from the constraints above: these constraints hold for some b if and only if

$$w^T(\hat{x}_+ - \hat{x}_-) \geq \kappa_\epsilon \left(\sqrt{w^T \Sigma_+ w} + \sqrt{w^T \Sigma_- w} \right).$$

Let us minimize the error probability level ϵ. Since κ_ϵ is decreasing in ϵ, we would like to maximize κ_ϵ subject to the constraints above. Exploiting homogeneity, we can always require $w^T(\hat{x}_+ - \hat{x}_-) = 1$. We are led to the problem

$$p^* = 1/\kappa^* := \min_w \; \sqrt{w^T \Sigma_+ w} + \sqrt{w^T \Sigma_- w} \;:\; w^T(\hat{x}_+ - \hat{x}_-) = 1.$$

This is an SOCP, which is strictly feasible (in the sense of weak Slater's condition), since $\hat{x}_+ \neq \hat{x}_-$. Hence strong duality holds.

The dual can be obtained from the minimax expression

$$p^* = \min_w \max_{\nu,\, \|u_\pm\|_2 \leq 1} u_+^T \Sigma_+^{1/2} w + u_-^T \Sigma_-^{1/2} w + \nu(1 - w^T(\hat{x}_+ - \hat{x}_-)).$$

Exchanging the min and max yields

$$p^* = d^* = \max_{\nu,\, \|u_\pm\|_2 \leq 1} \nu \;:\; u_+^T \Sigma_+^{1/2} + u_-^T \Sigma_-^{1/2} = \nu(\hat{x}_+ - \hat{x}_-).$$

We observe that $\nu \neq 0$ at the optimum, otherwise we would have $p^* = 0$, which is clearly impossible when $\hat{x}_+ \neq \hat{x}_-$. We then set $\kappa = 1/\nu$, scale the variables $u_\pm$ by r, and change u_+ into its opposite. This leads to the dual

$$\kappa^* := \min_{\|u_\pm\|_2 \leq \kappa} \kappa \;:\; \hat{x}_+ + \Sigma_+^{1/2} u_+ = \hat{x}_- + \Sigma_-^{1/2} u_-.$$

The geometric interpretation is as follows. Consider the ellipsoids $\mathcal{E}_\pm(\kappa) := \{\hat{x}_\pm + \Sigma_\pm^{1/2} u \,:\, \|u\|_2 \leq \kappa\}$. The constraints in the dual problem above state that these two ellipsoids intersect. The dual problem amounts to finding the smallest r for which the two ellipsoids $\mathcal{E}_\pm(\kappa)$ intersect. The optimal value of κ is then κ^*, and the corresponding optimal value of the error probability level is $\epsilon^* = \Phi(-\kappa^*)$. It can be shown that the optimal separating hyperplane corresponds to the common tangent at the intersection point.

Largest eigenvalue problem. Let us use the KKT conditions to prove that, for any given matrix $A \in \mathcal{S}^n$:

$$\max_x \{x^T A x \;:\; x^T x = 1\} = \max_{X \succeq 0,\, \mathbf{Tr}\, X = 1} \mathbf{Tr}\, AX = \lambda_{\max}(A),$$

where $\lambda_{\max}$ denotes the largest singular value.

Duality theory for SDP immediately tells us that the second equality holds. Indeed, the SDP

$$p^* = \max_X \ \mathbf{Tr}\, AX \ : \ X \succeq 0, \ \ \mathbf{Tr}\, X = 1 \tag{4.25}$$

admits the following dual:

$$p^* \leq d^* := \min_t \ t \ : \ tI \succeq A.$$

Using the eigenvalue decomposition of A, it is easy to show that $d^* = \lambda_{\max}(A)$.

It remains to prove the first equality. We observe that

$$\max_x \{x^T Ax \ : \ x^T x = 1\} = \max_X \ \mathbf{Tr}\, AX \ : \ X \succeq 0, \ \mathbf{Tr}\, X = 1, \ \ \mathbf{rank}(X) = 1.$$

(To see this, set $X = xx^T$.) Thus we need to show that at optimum, the rank of the primal variable X in (4.32) is one.

The pair of primal and dual problems are both strictly feasible, hence the KKT condition theorem applies, and both problems are attained by some primal-dual pair (X, t), which satisfies the KKT conditions. These are $X \succeq 0$, $tI \succeq A$, and $(tI - A)X = 0$. The last condition proves that any non-zero column x of X satisfies $(tI - A)x = 0$ (in other words, x is an eigenvector associated with the largest eigenvalue). Let us normalize x so that $\|x\|_2 = 1$, so that $\mathbf{Tr}\, xx^T = 1$. We have $(tI - A)xx^T = 0$, which proves that the feasible primal variable $X^* = xx^T \succeq 0$, $\mathbf{Tr}\, X^* = 1$, is feasible and optimal for the primal problem (4.32). Since X^* has rank one, our first equality is proved.

4.8 Optimality Conditions

4.8.1 Complementary slackness

We consider a primal convex optimization problem (without equality constraints for simplicity):

$$p^* := \min_x \ f_0(x) \ : \ f_i(x) \leq 0, \ \ i = 1, \ldots, m,$$

and its dual

$$p^* \geq d^* := \max_\lambda \ g(\lambda),$$

where g is the dual function

$$g(\lambda) := \min_x \ \mathcal{L}(x, \lambda) \left(:= f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) \right).$$

We assume that the duality gap is zero: $p^* = d^*$, and that both primal and dual values are attained, by a primal-dual pair (x^*, λ^*). We have

$$p^* = f_0(x^*) = d^* = g(\lambda^*) = \min_x \ \mathcal{L}(x, \lambda^*) \leq \mathcal{L}(x^*, \lambda^*) \leq f_0(x^*) = p^*. \tag{4.26}$$

Thus, equalities hold in the above.

This implies that x^* minimizes the function $\mathcal{L}(\cdot, \lambda^*)$:

$$x^* \in \arg\min_x \mathcal{L}(x, \lambda^*).$$

If the functions $f_0, \ldots, f_m$ are differentiable, the above implies

$$\nabla_x \mathcal{L}(x, \lambda^*)|_{x=x^*} := \nabla f_0(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0.$$

In addition, the equalities in (4.26) imply

$$\sum_{i=1}^m \lambda_i^* f_i(x^*) = 0.$$

Since $\lambda^* \geq 0$, $f_i(x^*) \leq 0$, $i = 1, \ldots, m$, the above is equivalent to the *complementary slackness condition:*

$$\lambda_i^* f_i(x^*) = 0, \quad i = 1, \ldots, m.$$

4.8.2 KKT optimality conditions

Assume that the functions $f_0, \ldots, f_m$ are differentiable.

Consider the so-called KKT[3] conditions on a primal-dual pair (x^*, λ^*).

$$\begin{array}{rr} f_i(x^*) \leq 0, \quad i = 1, \ldots, m & \text{(primal feasibility)}, \\ \lambda^* \geq 0 & \text{(dual feasibility)}, \\ \lambda_i^* f_i(x^*) = 0, \quad i = 1, \ldots, m & \text{(complementary slackness)}, \\ \nabla_x \mathcal{L}(x, \lambda^*)|_{x=x^*} = 0 & \text{(Lagrangian stationarity)}. \end{array} \tag{4.27}$$

The previous development shows that for any problem (convex or not) for which strong duality holds, and primal and dual values are attained, the KKT conditions (4.27) are *necessary* for a primal-dual pair (x^*, λ^*) to be optimal.

If, in addition the problem is convex, then the conditions are also *sufficient.* To see this, note that the first two conditions imply that x^*, λ^* are feasible for the primal and dual problems, respectively. Since $\mathcal{L}(\cdot, \lambda^*)$ is convex, the fourth condition (which we called Lagrangian stationarity) states that x^* is a minimizer of $\mathcal{L}(\cdot, \lambda^*)$, hence

$$g(\lambda^*) = \min_x \mathcal{L}(x, \lambda^*) = \mathcal{L}(x^*, \lambda^*) = f_0(x^*),$$

where we used the third condition (complementary slackness) for the last equality. The above proves that the primal-dual pair (x^*, λ^*) is optimal, since the corresponding gap is zero.

4.8.3 Primal solutions from dual variables

Assume that the problem has a zero duality gap, withdual values attained. Now assume that λ^* is optimal for the dual problem, and assume that the minimization problem

$$\min_x \mathcal{L}(x, \lambda^*).$$

has a unique solution $x(\lambda^*)$ that is feasible for the primal problem. Then, $x(\lambda^*)$ is optimal. Indeed, the fourth KKT condition (Lagrange stationarity) states that any optimal primal point minimizes the partial Lagrangian $\mathcal{L}(\cdot, \lambda^*)$, so it must be equal to the unique minimizer $x(\lambda^*)$.

This allows to compute the primal solution when a dual solution is known, by solving the above problem.

[3]The acronym comes from the names Karush, Kuhn and Tucker, researchers in optimization around 1940-1960.

4.9 Conic Duality

4.9.1 Conic problem and its dual

The conic optimization problem in standard equality form is:

$$p^* := \min_x \; c^T x \;:\; Ax = b, \;\; x \in \mathcal{K}.$$

where $\mathcal{K}$ is a proper cone, for example a direct product of cones that are one of the three types: positive orthant, second-order cone, or semidefinite cone. Let $\mathcal{K}^*$ be the cone dual $\mathcal{K}$, which we define as

$$\mathcal{K}^* := \{\lambda \;:\; \forall x \in \mathcal{K}, \;\; \lambda^T x \geq 0\}. \tag{4.28}$$

All cones we mentioned (positive orthant, second-order cone, or semidefinite cone), are self-dual, in the sense that $\mathcal{K}^* = \mathcal{K}$.

The Lagrangian of the problem is given by

$$\mathcal{L}(x, \lambda, y) = c^T x + y^T(b - Ax) - \lambda^T x \tag{4.29}$$

The last term is added to take account of the constraint $x \in \mathcal{K}$. From the very definition of the dual cone:

$$\max_{\lambda \in \mathcal{K}^*} \; -\lambda^T x = \begin{cases} 0 & \text{if } x \in \mathcal{K}, \\ +\infty & \text{otherwise.} \end{cases}$$

Thus, we have

$$\begin{aligned} p^* &= \min_x \max_{y, \lambda \in \mathcal{K}^*} \mathcal{L}(x, \lambda, y) \\ &= \min_x \max_{y, \lambda \in \mathcal{K}^*} c^T x + y^T(b - Ax) - \lambda^T x \\ &\geq d^* := \max_{y, \lambda \in \mathcal{K}^*} \; g(\lambda, y) \end{aligned} \tag{4.30}$$

where

$$g(\lambda, y) = \min_x \; c^T x + y^T(b - Ax) - \lambda^T x = \begin{cases} y^T b & \text{if } c - A^T y - \lambda = 0, \\ -\infty & \text{otherwise} \end{cases}$$

The dual for the problem is:

$$d^* = \max \; y^T b \;:\; c - A^T y - \lambda = 0, \;\; \lambda \in \mathcal{K}^*.$$

Eliminating λ, we can simplify the dual as:

$$d^* = \max \; y^T b \;:\; c - A^T y \in \mathcal{K}^*.$$

4.9.2 Conditions for strong duality

We now summarize the results stated in past lectures. Strong duality hold when either:

- The primal is strictly feasible, i.e. $\exists x: \;\; Ax = b, \;\; x \in \mathbf{int}(\mathcal{K})$. This also implies that the dual problem is attained.
- The dual is strictly feasible, i.e. $\exists y: \; c - A^T y \in \mathbf{int}(\mathcal{K}^*)$. This also implies that the primal problem is attained.
- If both the primal and dual are strictly feasible then both are attained (and $p^* = d^*$).

4.9.3 KKT conditions for conic problems

Assume $p^* = d^*$ and both the primal and dual are attained by some primal-dual triplet (x^*, λ^*, y^*). Then,

$$\begin{aligned} p^* = c^T x^* = d^* &= g(\lambda^*, y^*) \\ &= \min_x \mathcal{L}(x, \lambda^*, y^*) \\ &\leq \mathcal{L}(x^*, \lambda^*, y^*) \\ &= c^T x^* - \lambda^{*T} x^* + y^{*T}(b - Ax^*) \\ &\leq c^T x^* = p^*. \end{aligned} \tag{4.31}$$

The last term in the fourth line is equal to zero which implies $\lambda^{*T} x^* = 0$.

Thus the KKT conditions are:

- $x \in \mathcal{K}$, $Ax = b$,
- $\lambda \in \mathcal{K}^*$,
- $\lambda^T x = 0$,
- $c - A^T y - \lambda = 0$, that is, $\nabla_x \mathcal{L}(x, \lambda, y) = 0$.

Eliminating λ from the above allows us to get rid of the Lagrangian stationarity condition, and gives us the following theorem.

Theorem 13 (KKT conditions for conic problems) *The conic problem*

$$p^* := \min_x \; c^T x \;\; : \;\; Ax = b, \;\; x \in \mathcal{K}.$$

admits the dual bound $p^* \geq d^*$, where

$$d^* = \max \; y^T b \;\; : \;\; c - A^T y \in \mathcal{K}^*.$$

If both problems are strictly feasible, then the duality gap is zero: $p^* = d^*$, and both values are attained. Then, a pair (x, y) is primal-dual optimal if and only if the KKT conditions

1. Primal feasibility: $x \in \mathcal{K}$, $Ax = b$,
2. Dual feasibility: $c - A^T y \in \mathcal{K}^*$,
3. Complementary slackness: $(c - A^T y)^T x = 0$,

hold.

4.9.4 KKT conditions for SDP

Consider the SDP

$$p^* = \min_x \; c^T x \;\; : \;\; F(x) := F_0 + \sum_{i=1}^m x_i F_i \succeq 0,$$

with $c \in \mathbf{R}^n$, $F_i \in \mathcal{S}^n$, $i = 0, \ldots, m$. The Lagrangian is

$$\mathcal{L}(x, Y) = c^T x - \mathbf{Tr}\, F(x) Y,$$

and the dual problem reads

$$d^* = \max_{Y \succeq 0} \min_x \; \mathcal{L}(x, Y) = \max_Y \; -\mathbf{Tr}\, F_0 Y \;\; : \;\; \mathbf{Tr}\, F_i Y = c_i, \;\; i = 1, \ldots, n, \; ; Y \succeq 0.$$

The KKT conditions for optimality are as follows:

1. $F(x) \succeq 0$,
2. $Y \succeq 0$, $\mathbf{Tr}(F_i Y) = c_i$, $i = 1, \ldots, n$,
3. $\mathbf{Tr}(F(x)Y) = 0$.

The last condition can be expressed as $F(x)Y = 0$, according to the following result: Let $F, Y \in \mathcal{S}^n$. If $F \succeq 0$ and $Y \succeq 0$ then $\mathbf{Tr}(FY) = 0$ is equivalent to $FY = 0$.

Proof: Let $Y^{1/2}$ be the square root of Y (the unique positive semi-definite solution to $Z^2 = Y$). We have $\mathbf{Tr}\, FY = \mathbf{Tr}\, \tilde{F} = 0$, where $\tilde{F} := Y^{1/2} F Y^{1/2}$. Since $F \succeq 0$, we have $\tilde{F} \succeq 0$. The trace of $\tilde{F}$ being zero then implies that $\tilde{F} = 0$.

Using the eigenvalue decomposition, we can reduce the problem to the case when Y is diagonal. Let us assume that

$$Y = \begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} F_{11} & F_{12} \\ F_{12}^T & F_{22} \end{pmatrix}$$

where $\Lambda \succ 0$ is diagonal, and contains the eigenvalues of Y, and the matrix F_{11} is of the same size as Λ (which is equal to the rank of Y). The condition $\tilde{F} = Y^{1/2} F Y^{1/2} = 0$ expresses as

$$0 = \begin{pmatrix} \Lambda^{1/2} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} F_{11} & F_{12} \\ F_{12}^T & F_{22} \end{pmatrix} \begin{pmatrix} \Lambda^{1/2} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \Lambda^{1/2} F_{11} \Lambda^{1/2} & 0 \\ 0 & 0 \end{pmatrix}.$$

Since $\Lambda \succ 0$, we obtain $F_{11} = 0$. But $F \succeq 0$ then implies $F_{12} = 0$ (use Schur complements), thus

$$FY = \begin{pmatrix} 0 & 0 \\ 0 & F_{22} \end{pmatrix} \begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix} = 0,$$

as claimed. Thus the last KKT condition can be written as $F(x)Y = 0$.

Theorem 14 (KKT conditions for SDP) *The SDP*

$$p^* := \min_x \; c^T x \;\; : \;\; F(x) := F_0 + \sum_{i=1}^m x_i F_i \succeq 0$$

admits the dual bound $p^* \geq d^*$, where

$$d^* = \max_Y \; -\mathbf{Tr}\, F_0 Y \;\; : \;\; \mathbf{Tr}\, F_i Y = c_i, \;\; i = 1, \ldots, n, \;\; Y \succeq 0.$$

If both problems are strictly feasible, then the duality gap is zero: $p^* = d^*$, and both values are attained. Then, a pair (x, Y) is primal-dual optimal if and only if the KKT conditions

1. Primal feasibility: $F(x) \succeq 0$,

2. Dual feasibility: $\mathbf{Tr}\, F_i Y = c_i, \;\; i = 1, \ldots, n, \;\; Y \succeq 0$,

3. Complementary slackness: $F(x)Y = 0$,

hold.

Recall LP duality for a problem of the from $\min_x \{c^T x : Ax \leq b\}$ with dual variables y has the KKT conditions were $\forall i : y_i (b - Ax)_i = 0$. This can be compactly written as $\mathbf{diag}(y)\, \mathbf{diag}(b - Ax) = 0$ which is similar to the KKT conditions for SDP (with $Y = \mathbf{diag}(y)$). This should come as no surprise as SDP problems include LP problems as a special case.

4.9.5 Examples

Largest eigenvalue problem. Let us use the KKT conditions to prove that, for any given matrix $A \in \mathcal{S}^n$:

$$\max_x \{x^T A x \ : \ x^T x = 1\} = \max_{X \succeq 0,\ \mathbf{Tr}\, X = 1} \mathbf{Tr}\, AX = \lambda_{\max}(A),$$

where $\lambda_{\max}$ denotes the largest singular value.

Duality theory for SDP immediately tells us that the second equality holds. Indeed, the SDP

$$p^* = \max_X \mathbf{Tr}\, AX \ : \ X \succeq 0, \ \ \mathbf{Tr}\, X = 1 \tag{4.32}$$

admits the following dual:

$$p^* \leq d^* := \min_t \ t \ : \ tI \succeq A.$$

Using the eigenvalue decomposition of A, it is easy to show that $d^* = \lambda_{\max}(A)$.

It remains to prove the first equality. We observe that

$$\max_x \{x^T A x \ : \ x^T x = 1\} = \max_X \mathbf{Tr}\, AX \ : \ X \succeq 0,\ \mathbf{Tr}\, X = 1, \ \ \mathbf{rank}(X) = 1.$$

(To see this, set $X = xx^T$.) Thus we need to show that at optimum, the rank of the primal variable X in (4.32) is one.

The pair of primal and dual problems are both strictly feasible, hence the KKT condition theorem applies, and both problems are attained by some primal-dual pair (X, t), which satisfies the KKT conditions. These are $X \succeq 0$, $tI \succeq A$, and $(tI - A)X = 0$. The last condition proves that any non-zero column x of X satisfies $(tI - A)x = 0$ (in other words, x is an eigenvector associated with the largest eigenvalue). Let us normalize x so that $\|x\|_2 = 1$, so that $\mathbf{Tr}\, xx^T = 1$. We have $(tI - A)xx^T = 0$, which proves that the feasible primal variable $X^* = xx^T \succeq 0$, $\mathbf{Tr}\, X^* = 1$, is feasible and optimal for the primal problem (4.32). Since X^* has rank one, our first equality is proved.

4.10 Sensitivity Analysis

The material of this section is entirely contained in section §5.6 of BV.

Chapter 5

Algorithms

5.1 Interior-Point Methods

The material of this lecture is entirely contained in section §5.6 of BV.

5.2 First-Order Methods (I)

5.2.1 Motivation

Interior point methods have complexity of $C \cdot O(\log(\frac{1}{\epsilon}))$, where C is dependent on problem size. These methods are second-order methods, as they require the Hessian to be evaluated, and the corresponding KKT system of the form $Hx = g$ needs to be solved at each Newton step. Thus, C can be quite high for large-scale problems. On the other hand, the dependence on the accuracy ϵ is very good.

The motivation for first-order methods is:

- a much better complexity per iteration;
- at the price of a worse dependence on the precision ϵ, anywhere from $\log(1/\epsilon)$ to $1/\epsilon^2$.

Many of these methods hinge on the notion of subgradient, which generalizes to arbitrary convex functions the notion of gradient for differentiable functions.

5.2.2 Subgradients

5.2.3 Definition

Let $f : \mathbf{R}^n \to \mathbf{R}$ be a convex function. The vector $g \in \mathbf{R}^n$ is a *subgradient* of f at x if

$$\forall\, y \;:\; f(y) \geq f(x) + g^T(y - x).$$

The *subdifferential* of f at x, denoted $\partial f(x)$, is the set of such subgradients at x.

- $\partial f(x)$ is convex, closed, never empty on the relative interior[1] of its domain.

[1] The relative interior of a set is the interior of the set, relative to the smallest affine subspace that contains it.

- if f is differentiable at x, then the subdifferential is a singleton: $\partial f(x) = \{\nabla f(x)\}$.

For example, consider $f(x) = |x|$ for $x \in \mathbf{R}$. We have

$$\partial f(x) = \begin{cases} \{-1\} & \text{if } x < 0, \\ [-1, 1] & \text{if } x = 0, \\ \{+1\} & \text{if } x > 0. \end{cases}$$

5.2.4 Constructing subgradients

Weak rule for point-wise supremum: if f_α are differentiable and convex functions that depend on a parameter $\alpha \in \mathcal{A}$, with $\mathcal{A}$ an arbitrary set, then

$$f(x) = \sup_{\alpha \in \mathcal{A}} f_\alpha(x)$$

is possibly non-differentiable but convex. If β is such that $f(x) = f_\beta(x)$, then a subgradient of f at x is simply any element in $\partial f_\beta(x)$.

Example: maximum eigenvalue. For $X = X^T \in \mathbf{R}^{n \times n}$, define $f(X) = \lambda_{\max}(X)$ to be the largest eigenvalue of X (f is real valued since X is symmetric). A subgradient of f at X can be found using the following variational (that is, optimization-based) representation of $f(X)$:

$$f(X) = \max_{y \,:\, \|y\|_2 = 1} y^T X y.$$

Any unit-norm eigenvector $y_{\max}$ of X corresponding to the largest eigenvalue achieves the maximum in the above. Hence, by the weak rule above, a subgradient of f at X is given by a gradient of the function $X \to y_{\max}^T X y_{\max}$, which is $y_{\max} y_{\max}^T$.

Strong rule for pointwise maximum. If a function is defined as the pointwise maximum of convex functions, we can compute the whole sub-differential (and not just one subgradient via the weak rule). The strong rule is: if $f = \max_i f_i$, then

$$\partial f(x) = \mathbf{Co} \bigcup_i \{\partial f_i(x) \ : \ f_i(x) = f(x)\}.$$

That is, the subdifferential is the convex hull of union of subdifferentials of active functions at x.

Minimization rule. Assume that the function f is given as the result of an optimization problem:

$$f(y) = \min_x \ f_0(x) \ : \ f_i(x) \leq y_i, \ \ i = 1, \ldots, m.$$

The problem has a Lagrangian of the form

$$\mathcal{L}(x, \lambda) = f_0(x) + \sum_{i=1}^m \lambda_i (f_i(x) - y_i),$$

with the dual variable $\lambda_i \geq 0$, $i = 1, \ldots, m$. Assume that the primal problem is strictly feasible. Then, an optimal dual variable λ^* exists, and we have

$$h(y) = \min_x \ f_0(x) + \sum_{i=1}^m \lambda_i^* (f_i(x) - y_i). \tag{5.1}$$

We have

$$\begin{aligned}
h(z) &= \min_x \max_{\lambda \geq 0} f_0(x) + \sum_{i=1}^m \lambda_i (f_i(x) - z_i) \text{ [unconstrained representation]} \\
&\geq \max_{\lambda \geq 0} \min_x f_0(x) + \sum_{i=1}^m \lambda_i (f_i(x) - z_i) \text{ [weak duality for } h(z)] \\
&\geq \min_x f_0(x) + \sum_{i=1}^m \lambda_i^* (f_i(x) - z_i) \text{ [choose } \lambda = \lambda^* \text{ in the above]} \\
&= \min_x f_0(x) + \sum_{i=1}^m \lambda_i^* (f_i(x) - y_i) + (y - z)^T \lambda^* \\
&= h(y) + g^T (z - y), \text{ [eqn. (5.1)]}
\end{aligned}$$

where $g = -\lambda^*$. Hence $-\lambda^*$ is a subgradient of h at y.

5.2.5 Optimality

We consider the unconstrained minimization problem

$$\min_x f(x),$$

where f is convex.

Differentiable case. If f is differentiable, then the condition for optimality is just $\nabla f(x) = 0$.

Non-differentiable case. If f is not differentiable, but convex, then the condition for x to be optimal is

$$\forall\, y \ : \ f(y) \geq f(x).$$

This can be written equivalently as $0 \in \partial f(x)$, since by the strong rule given before:

$$0 \in \partial f(x) \Leftrightarrow \exists\, x \ : \ \forall\, y, \ \ f(y) \geq f(x) + 0^T (y - x) = f(x). \tag{5.2}$$

Example: piece-wise linear minimization

We show in this example how the classical LP duality and this new condition for optimality are the same.

Let

$$f(x) = \max_i (a_i^T x + b_i)$$

then

$$\min \ \ f(x) = \min \ t \ : \ t \geq a_i^T x + b_i, \ \ i = 1, \ldots, m.$$

We get the dual to be:

$$\max_\lambda b^T \lambda : \ : \ \lambda \geq 0, \ \ A^T \lambda = 0, \ \ 1^T \lambda = 1.$$

The corresponding KKT conditions are

$$t \geq a_i^T x + b_i, \ \ \lambda \geq 0, \ ; \lambda_i = 0 \text{ if } t > a_i^T x + b_i, \ \ 0 = \sum \lambda_i a_i, \ \ \sum \lambda_i = 1.$$

The above can be written equivalently as $0 \in \partial f(x)$, since

$$\partial f(x) = \mathbf{Conv}\{a_i \ : \ f(x) = a_i^T x + b_i\} = \left\{ \sum_{i \in I(x)} \lambda_i a_i \ : \ \sum_{i \in I(x)} \lambda_i = 1 \right\},$$

where $I(x)$ is the set of indices i that achieve the maximum in $\max_i(a_i^T x + b_i)$.

5.2.6 Subgradient Methods

The subgradient method is a simple algorithm for minimizing a non-differentiable convex function. The subgradient method uses step lengths that are fixed ahead of time, instead of an exact or approximate line search as in the gradient method. Unlike the ordinary gradient method, the subgradient method is not a descent method; the function value can (and often does) increase. The subgradient method is far slower than Newton's method, but is much simpler and can be applied to a far wider variety of problems.

5.2.7 Unconstrained case

Suppose $f : \mathbf{R} \to \mathbf{R}^n$ is convex, and we seek to solve the problem

$$p^* = \min_x \ f(x).$$

To minimize f, the subgradient method uses only subgradient information at every step.

Algorithm:

$$x^{(k+1)} = x^{(k)} - \alpha_k g^{(k)}, \ \ g^{(k)} \in \partial f(x^{(k)}).$$

Here, $x^{(k)}$ is the k-th iterate, $g^{(k)}$ is any subgradient of f at $x^{(k)}$, and $\alpha_k > 0$ is the k-th step size. Thus, at each iteration of the subgradient method, we take a step in the direction of a negative subgradient. Since this is not a descent method so we must keep track of the best solution seen so far, via the values

$$p_{\text{best}}^{(k)} = \min_{0 \le i \le k} \ f(x^{(k)}).$$

We will also denote $\overline{p} = \lim_{k \to \infty} p_{\text{best}}^{(k)}$.

Assumptions: The assumptions for subgradient methods are as follows:

- p^* is finite and attained
- There is a constant G such that, for every x, and every $g \in \partial f(x)$, we have $\|g\|_2 \le G$.
- There is a large enough constant R such that $R \ge ||x^{(1)} - x^*||_2$.

Step-size rules: Several different step size rules can be used.

- Constant step size: $\alpha_k = \alpha \ \ \forall k$
- Constant step length: $\alpha_k = \gamma / \|g^{(k)}\|_2$. This implies that $\|x^{(k+1)} - x^k\|_2 \le \gamma$ for every k.

Convergence analysis. Let $x^\star$ be any minimizer of f. We have

$$\begin{aligned}\|x^{(k+1)} - x^\star\|_2^2 &= \|x^{(k)} - \alpha_k g^{(k)} - x^\star\|_2^2 \\ &= \|x^{)k)} - x^\star\|_2^2 - 2\alpha_k g^{(k)T}(x^{(k)} - x^\star) + \alpha_k^2\|g^{(k)}\|_2^2 \\ &\leq \|x^{(k)} - x^\star\|_2^2 - 2\alpha_k(f(x^{(k)} - p^\star) + \alpha_k^2\|g^{(k)}\|_2^2,\end{aligned}$$

where we have used

$$p^\star = f(x^\star) \geq f(x^{(k)}) + g^{(k)T}(x^\star - x^{(k)}).$$

Applying this recursively, we get

$$\begin{aligned}\|x^{(k+1)} - x^\star\|_2 &\leq \|x^{(1)} - x^\star\|_2^2 - 2\sum_{i=1}^k (f(x^{(i)} - p^\star) + \sum_{i=1}^k \alpha_i^2\|g^{(i)}\|_2^2 \\ &\leq R^2 - 2\sum_{i=1}^k (f(x^{(i)} - p^\star) + G^2\sum_{i=1}^k \alpha_i^2.\end{aligned}$$

By definition of $p_{\text{best}}^{(k)}$, we have

$$\sum_{i=1}^k (f(x^{(i)} - p^\star) \geq (p_{\text{best}}^{(k)} - p^\star)\left(\sum_{i=1}^k \alpha_i\right),$$

which yields

$$p_{\text{best}}^{(k)} - p^\star \leq \frac{R^2 + G^2\sum_{i=1}^k \alpha_i^2}{2\sum_{i=1}^k \alpha_i}. \tag{5.3}$$

This shows that

- For constant step sizes, the condition $G^2\alpha/2 \leq \epsilon$ guarantees that $\overline{p} - p^\star \leq G^2\alpha/2 \leq \epsilon$.
- For constant step length, the condition $G\gamma/2 \leq \epsilon$ guarantees that $\overline{p} - p^\star \leq G^2\alpha/2 \leq \epsilon$.

Complexity. The convergence analysis result (5.3) depends on the choice of the step size rule. Which rule is optimal for this bound? The problem of minimizing the upper bound in (5.3) is convex and symmetric (the function does not change when we exchange variables). Hence, the optimal α_i's are all equal at optimum, to a number $\alpha = (R/G)k^{-1/2}$. With this choice of step length rule, the number of iterations needed to achieve ϵ-sub-optimality, as predicted by the analysis, is $(RG)/\epsilon^2$. The dependence on ϵ is now $O(1/\epsilon^2)$, which is much worse than that delivered by interior-point methods (which have $O(\log(1/\epsilon))$). This is not surprising: sub-gradient methods apply to any convex problem (provided we are able to compute a sub-gradient), whereas IPMs only apply to specific convex problems.

Example. We seek to find a point in a given intersection of closed convex sets,

$$C = C_1 \cap \cdots \cap C_m \subseteq R^n.$$

We formulate this problem as minimizing f, where

$$f(x) = \max(\mathbf{dist}(x, C_1) \cdots \mathbf{dist}(x, C_m)).$$

Let P_j the projection operator on C_j, and let

$$f_j(x) = \mathbf{dist}(x, C_j)$$

be the corresponding distance function. Thus, $P_j(x)$ achieves the minimum value of f_j. Then, a subgradient of f_j at x is

$$g_j(x) = \frac{x - P_j(x)}{\|x - P_j(x)\|_2}.$$

5.3 First-Order Methods (II)

Reading assignment: Notes on decomposition methods by Stephen Boyd, Lin Xiao, Almir Mutapcic and Jacob Mattingley posted on bspace.

5.3.1 Constrained Case

We now consider the convex, inequality constrained problem:

$$\min\ f_0(x)\ :\ f_i(x) \leq 0,\ \ i = 1, \ldots, m.$$

It is sometimes useful to consider the problem in abstract form:

$$\min_x\ f(x)\ :\ x \in C,$$

where C is a convex set.

5.3.2 Projected Gradient Method

Algorithm. The projected subgradient method uses the iteration

$$x^{(k+1)} = P_C(x^{(k)} - \alpha(k)g(k))$$

where P_C is projection on C:

$$P_C(x) = \arg\min_{z \in C}\ \|x - z\|_2,$$

and $g(k)$ is any subgradient of f at $x(k)$.

Analysis. The convergence analysis is the same as the ordinary subgradient method for unconstrained problems. The reason for this is that projection does not increase Euclidean length.

Examples.

- *Equality-constrained problems.* Let $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$, with $m \leq n$, A full rank (hence, $AA^T \succ 0$). We consider the problem

  $$\min_x\ f(x)\ :\ Ax = b,$$

  where $A \in \mathbf{R}^{m \times n}$, with $m \leq n$, and A full (row) rank (that is, $AA^T \succ 0$). The projection on the affine space $\{x\ :\ Ax = b\}$ is

  $$P := I - A^T(AA^T)^{-1}A.$$

Indeed, the problem

$$\min_z \ \|z - x\|_2 \ : \ Ax = b$$

admits the unique solution $z = Px$ (check this!).

The subgradient iterations are thus given by:

$$x^{(k+1)} = x^{(k)} - \alpha^{(k)} P(g^{(k)}).$$

- *l_1-norm minimization.* For the problem

$$\min_x \ \|x\|_1 \ : \ Ax = b.$$

Let us derive a subgradient for the function $x \to \|x\|_1$. We have

$$\|x\|_1 = \max_{u \,:\, \|u\|_\infty \leq 1} u^T x,$$

hence, by the maximum rule for subgradients, we obtain that a subgradient for the l_1-norm function at x is $u^*(x)$, where $u^*(x)$ is any maximizer for the problem above. In particular, the choice $u^*(x) = \mathrm{sign}(x)$ works.

Thus, a subgradient of the objective is $g = \mathbf{sign}(x)$, so the projected subgradient update is

$$x^{(k+1)} = x^{(k)} - \alpha^{(k)} P(\mathbf{sign}(x^{(k)})).$$

5.3.3 Projected Subgradient for the Dual

The projected subgradient method may also be applied to the dual problem when

- Primal iterates are not feasible but become so as $k \to \infty$.
- Strong duality holds, and the dual function values converge to p^*.

We consider the problem

$$\min_x \ f_0(x) \ : \ f_i(x) \leq 0, \ \ i = 1, \ldots, m.$$

The dual problem is :

$$\max_{\lambda \geq 0} \ g(\lambda),$$

where

$$g(\lambda) = \min_x \ f_0(x) + \sum_{i=1}^m \lambda_i f_i(x).$$

The dual variable can be iterated as

$$\lambda^{(k+1)} = \max(\lambda^{(k)} - \alpha^{(k)} h^{(k)}, 0), \quad h^{(k)} \in \partial(-g(\lambda^{(k)})).$$

Note that, according to the subgradient construction rules,

$$h_i = -f_i(x^*(\lambda)), \ \ i = 1, \ldots, m,$$

where $x^*(\lambda) \in \mathrm{argmin}\,[f_0(x) + \sum \lambda_i f_i(x)]$.

Example. Let $P \succ 0$. Consider the problem

$$\min_x \ \frac{1}{2}x^T P x + q^T x, \ \ : \ x_i^2 \leq 1, \ \ i = 1, \ldots, m.$$

The Lagrangian is, with $D_\lambda := \mathbf{diag}(\lambda) \succeq 0$:

$$L(x, \lambda) = \frac{1}{2}x^T (P + 2D_\lambda)x - q^T x - 1^T \lambda.$$

A minimizer for the dual function is unique, and given by:

$$x^*(\lambda) := (P + 2D_\lambda)^{-1} q.$$

The iterates are of the form

$$\lambda^{(k+1)} = \max(\lambda^{(k)} - \alpha^{(k)} h^{(k)}, 0), \ \ h^{(k)} = 1 - (x^{(k)})^2, \ \ x^{(k)} = (P + 2D_{\lambda^{(k)}})^{-1} q.$$

5.3.4 Decomposition Methods

Decomposition methods are useful when attempting to solve large-scale (convex) optimization problems with a few coupling constraints or variables.

To illustrate some of the ideas, consider a problem of the form

$$p^* := \min_{x_1, x_2, y} \ f_1(x_1, y) + f_2(x_2, y),$$

where f_i's are both *jointly* convex. (Hence, the problem is convex.)

5.3.5 Primal decomposition

You can think of y as a *coupling variable,* which couples the behavior of the two terms. Primal decomposition is based on the observation that for fixed y, the problem decomposes as two problems involving independent variables. Precisely,

$$p^* = \min_y \ \phi_1(y) + \phi_2(y), \tag{5.4}$$

where ϕ_i's are defined as

$$\phi_i(y) = \min_x \ f_i(x, y), \ \ i = 1, 2.$$

Note that computing ϕ_i, $i = 1, 2$ amounts to solve two separate convex sub-problems (which can be processed on two separate computers). Note also that the function ϕ_i, $i = 1, 2$ are both convex, because of the joint convexity of the f_i's.

The "master" problem (5.4) can be solved by a sub-gradient method. All it requires is forming a subgradient of the objective function, which is of the form $g_1 + g_2$, where g_i is a subgradient of ϕ_i at y.

Consider the problem of finding a subgradient for the function

$$\phi(y) := \min_x \ f(x, y),$$

where f is a convex function of the variable (x, y). To do this, we assume that for every y, the solution to the above problem is attained by some optimal point $x^*(y)$. Since $x^*(y)$ is optimal, a subgradient of f at the point $z := (x^*(y), y)$ is of the form $(0, g(y))$. (For example, if f differentiable, the above means that

the partial derivative of f with respect to the first variable is zero at z.) Now, for every (x', y'), we have the subgradient inequality

$$f(x', y') \geq f(x^*(y), y) + \begin{pmatrix} 0 \\ g(y) \end{pmatrix}^T \begin{pmatrix} x' - x^*(y) \\ y' - y \end{pmatrix} = \phi(y) + g(y)^T(y' - y).$$

Since the left-hand side is independent of x', and the above is valid for every x', we can take the minimum of the right-hand side over x', and obtain

$$\phi(y') = \min_{x'} f(x', y') \geq \phi(y) + g(y)^T(y' - y),$$

which proves that $g(y)$ is a subgradient of ϕ at y.

5.3.6 Dual decomposition

In dual decomposition, we write the original problem as

$$p^* = \min_{x_1, x_2, y_1, y_2} f_1(x_1, y_1) + f_2(x_2, y_2) \ : \ y_1 = y_2.$$

Assuming strong duality, we can express the problem as

$$p^* = \max_{\nu} \min_{x_1, x_2, y_1, y_2} f_1(x_1, y_1) + f_2(x_2, y_2) + \nu^T(y_1 - y_2) = \max_{\nu} g_1(\nu) + g_2(\nu),$$

where

$$g_1(\nu) := \min_{x,y} f_1(x, y) + \nu^T y, \quad g_2(\nu) := \min_{x,y} f_2(x, y) - \nu^T y. \tag{5.5}$$

We can solve the above "master" problem using a subgradient method. To do this, we need to find the subgradients of the convex functions $-g_i$ at a given point ν. For a fixed ν, the problem of computing g_1, g_2 (and subgradients) can be solved separately.

After a change of sign, the problem boils down to the following: given a convex function of the form

$$h(\nu) = \max_{y} \nu^T y - f(x, y),$$

with f convex, find a subgradient of h at ν. We can apply the maximum rule for subgradients. Assuming that the above problem is attained at some variable $y^*(\nu)$, we obtain that $y^*(\nu)$ is a subgradient of h at ν.

The subgradient update will have the form

$$\nu_{k+1} = \nu_k - \alpha_k(y_2^*(\nu_k) - y_1^*(\nu_k)),$$

where $y_i^*(\nu)$ $(i = 1, 2)$ is any minimizer corresponding to the two separate problems (5.5). The above can be interpreted as a simple linear feedback rule, where the penalty parameters (ν_k) are updated according to how big the violation of the equality constraint $y_1 = y_2$ is.

Chapter 6

Applications

6.1 Moment Inequalities

Reading assignment: Section §7.4 of BV, and the Chapter 3 of the book on Robust Optimization[1]

6.1.1 Chance Linear Programming

See the Chapter mentioned above.

6.1.2 Bounds on Probabilities

Problem statement. We consider a random variable x with distribution π, which is only known to belong to a class of distributions Π, and seek a bound on the probability of a given set $\mathcal{C}$, that is, a lower bound on

$$\inf_{\pi \in \Pi} \mathbf{Prob}\, \mathcal{C}. \tag{6.1}$$

Alternatively, we may seek an upper bound on the quantity obtained by replacing "inf" with "sup" in the above. Both problem are equivalent, in the sense that replacing $\mathcal{C}$ by its complement in one problem leads to the other.

We assume that Π is a class of distributions with given mean and covariance matrix:

$$\Pi = \left\{ \pi \in \Pi_0 \ : \ \mathbf{E}_\pi x = \hat{x}, \ \ \mathbf{E}_\pi (x - \hat{x})(x - \hat{x})^T = \Sigma \right\},$$

where Π_0 is the set distributions on $\mathbf{R}^n$, $\hat{x} \in \mathbf{R}^n$, $\Sigma \in \mathcal{S}^n_{++}$ are given, and $\mathbf{E}_\pi$ denotes the expectation operator with respect to the distribution π.

Problems involving bounds on probabilities arise in many situations. For example, we may interested in *yield maximization,* which involves the function $Y(y) := \mathbf{Prob}(y + x \in S)$, where y is a vector of design parameters, x represents additive implementation noise, and S is a a subset of allowable designs.

Dual problem. We can formulate problem (6.1) as an infinite dimensional linear programming problem:

$$p^* := \inf_{\pi(\cdot) \geq 0} \int \mathbf{1}_{\mathcal{C}}(x) \pi(x) dx \ : \ \int \begin{pmatrix} x \\ 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix}^T \pi(x) dx = \Gamma, \tag{6.2}$$

[1] A. Ben Tal, L. El Ghaoui, A. Nemirovski, *Robust Optimization,* 2009.

where $\mathbf{1}_{\mathcal{C}}$ is the location function of $\mathcal{C}$ (with value 1 on $\mathcal{C}$, and 0 elsewhere), and

$$\Gamma = \begin{pmatrix} \Sigma + xx^T & x \\ x^T & 1 \end{pmatrix} \succ 0.$$

The problem is linear in the sense that it involves an objective that is linear (in the variable π), affine equality constraints, and sign constraints. Of course, this is not an LP in the classical sense, as the variable is infinite-dimensional.

Using duality, we can transform the problem into one with infinitely many constraints, and finitely many variables. To do this, we first obtain a weak duality result, using the Lagrange functional[2]

$$\mathcal{L}(\pi, M) = \inf_{\pi(\cdot)\geq 0} \int \mathbf{1}_{\mathcal{C}}(x)\pi(x)dx + \langle M, \Gamma - \int \begin{pmatrix} x \\ 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix}^T \pi(x)dx \rangle.$$

We check that this Lagrangian "works", that is, we have the minimax representation

$$p^* := \min_{\pi(\cdot)\geq 0} \max_{M \in \mathcal{S}^n} \mathcal{L}(\pi, M).$$

By weak duality, we have $d^* \leq p^*$, with

$$d^* := \max_{M \in \mathcal{S}^n} \min_{\pi(\cdot)\geq 0} \mathcal{L}(\pi, M).$$

The dual function is

$$g(M) := \min_{\pi(\cdot)\geq 0} \mathcal{L}(\pi, M) = \langle M, \Gamma \rangle + \min_{\pi(\cdot)\geq 0} \langle \pi, \mathbf{1}_{\mathcal{C}} - q_M \rangle,$$

where q_M is the quadratic function with values

$$q_M(x) := \langle M, \begin{pmatrix} x \\ 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix}^T \rangle = \begin{pmatrix} x \\ 1 \end{pmatrix}^T M \begin{pmatrix} x \\ 1 \end{pmatrix},$$

and we define the scalar product between two measures π, h as

$$\langle \pi, h \rangle := \int \pi(x)h(x)dx.$$

It is easy to show that, for any function h:

$$\min_{\pi(\cdot)\geq 0} \langle \pi, h \rangle = \begin{cases} 0 & \text{if } h(x) \geq 0 \text{ for every } x \in \mathbf{R}^n, \\ -\infty & \text{otherwise.} \end{cases}$$

We obtain

$$g(M) = \begin{cases} \langle M, \Gamma \rangle & \text{if } \mathbf{1}_{\mathcal{C}}(x) \geq q_M(x) \text{ for every } x \in \mathbf{R}^n, \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem reads

$$\sup_{M=M^T} g(M) = \sup_{M=M^T} \langle M, \Gamma \rangle \ : \quad \begin{array}{l} \forall x \in \mathbf{R}^n, q_M(x) \leq 1, \\ \forall\, x \notin \mathcal{C}, \ \ q_M(x) \leq 0. \end{array}$$

The first constraint is equivalent to the semidefinite constraint $M \preceq J$, where J is a matrix with all zeros, except a 1 in the lower-right element.

[2]We say "functional" as the Lagrangian's input variables includes π, which is a function, more precisely a measure.

Further reductions. In some cases, the dual problem can be expressed exactly as a semidefinite program. Consider the case when $\mathcal{C}$ is defined by a *single* (possibly non-convex) quadratic inequality:

$$\mathcal{C} = \left\{ x \ : \ q(x) := \begin{pmatrix} x \\ 1 \end{pmatrix}^T Q \begin{pmatrix} x \\ 1 \end{pmatrix} > 0 \right\},$$

with $Q = Q^T$ given.

Then, using the $\mathcal{S}$-lemma (see BV, §B.2) the condition

$$\forall\, x, \ \ q(x) \le 0, \ \ q_M(x) \le 0$$

is equivalent to the existence of $\tau \ge 0$ such that $M \preceq \tau Q$.

The dual problem now reads

$$d^* = \sup_{M = M^T,\, \tau \ge 0} \langle M, \Gamma \rangle \ : \ J \succeq M, \ \ M \preceq \tau Q,$$

which is an SDP. It turns out that it can be further simplified greatly, to a single-variable convex problem. Precisely, we have

$$1 - d^* = \min_{\tau \ge 0} \lambda_{\max}[(J - \tau Q)_+],$$

where X_+ is the matrix obtained from the symmetric matrix X by replacing the negative eigenvalues by 0.

Strong duality. It can be shown that if $\Gamma \succ 0$, then strong duality holds.

Chebyschev and Markov inequalities. Chebyschev and Markov inequalities can be derived from the above, as special cases.