EE364a Review Session 5

announcements:

- homeworks 1 and 2 graded
- homework 4 solutions (check solution to additional problem 1)
- scpd phone-in office hours: tuesdays 6-7pm (650-723-1156)

Complementary slackness

consider the (not necessarily convex) optimization problem,

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p. \end{array}$$

let $x^\star$, and $(\lambda^\star, \nu^\star)$ be primal and dual optimal points, and suppose strong duality holds ($i.e.$, $p^\star = d^\star$). this means,

$$\begin{aligned} f_0(x^\star) &= g(\lambda^\star, \nu^\star) \\ &= \inf_x \left(f_0(x) + \sum_{i=1}^m \lambda_i^\star f_i(x) + \sum_{i=1}^p \nu_i^\star h_i(x) \right) \\ &\leq f_0(x^\star) + \sum_{i=1}^m \lambda_i^\star f_i(x^\star) + \sum_{i=1}^p \nu_i^\star h_i(x^\star) \\ &\leq f_0(x^\star). \end{aligned}$$

this result has several important conclusions.

- $x^\star$ minimizes $L(x, \lambda^\star, \nu^\star)$ over x (although note that $L(x, \lambda^\star, \nu^\star)$ can have other minimizers). we'll see an application of this soon.

- another conclusion is that

$$\sum_{i=1}^{m} \lambda_i^\star f_i(x^\star) = 0.$$

what does this mean?
answer.

$$\lambda^\star f_i(x^\star) = 0, \quad i = 1, \ldots, m.$$

or, equivalently,

$$\lambda_i^\star > 0 \Rightarrow f_i(x^\star) = 0,$$

or,

$$f_i(x^\star) < 0 \Rightarrow \lambda_i^\star = 0.$$

Example: minimizing a linear function over a rectangle

let's take a look at a familiar problem,

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & 0 \preceq x \preceq \mathbf{1}. \end{array}$$

- what is the solution?
  answer. if $c_i > 0$, then $x_i^\star = 0$, if $c_i < 0$ then $x_i^\star = 1$, and if $c_i = 0$, then $0 \leq x_i^\star \leq 1$.

- we can look at this from another point of view. let us associate dual variables λ with the constraint $0 \preceq x$, and dual variables μ with the constraint $x \preceq \mathbf{1}$. the Lagrangian is

$$L(x, \lambda, \mu) = c^T x - \lambda^T x + \mu^T (x - \mathbf{1}).$$

  this is bounded if and only if $c = \lambda - \mu$. any dual optimal point $(\lambda^\star, \mu^\star)$ must satisfy $c = \lambda^\star - \mu^\star$.

- suppose $c_i > 0$, what can we say about $\lambda_i^\star$ $x_i^\star$ and $\mu_i^\star$?
  answer. $\lambda_i^\star > 0$, so by complementary slackness $x_i^\star = 0$, and therefore $\mu_i^\star = 0$.

- suppose $c_i < 0$, what can we say about $\lambda_i^\star$ $x_i^\star$ and $\mu_i^\star$?
  answer. $\mu_i^\star > 0$, so by complementary slackness $x_i^\star = 1$, and therefore $\lambda_i^\star = 0$.

- what about when $c_i = 0$?
  answer. then $\lambda_i^\star - \mu_i^\star = 0$. but we cannot have both $\lambda_i^\star > 0$ and $\mu_i^\star > 0$, since that would imply both $x_i = 0$ and $x_i = 1$. as a result we must have $\lambda_i^\star = 0$ and $\mu_i^\star = 0$.

- can we write down $\lambda^\star$ and $\mu^\star$?
  answer. yes, $\lambda^\star$ is the positive part of c, and $\mu^\star$ is the negative part of c. *i.e.*, $\lambda^\star = \max(c, 0)$, $\mu^\star = \min(c, 0)$.

- the Lagrange multipliers λ and ν are also a measure of how active a constraint is at the optimum. for example, if we relax the first constraint by $-u_1 \leq x_1$. then

$$\lambda_1^\star = -\left.\frac{\partial p^\star}{\partial u_1}\right|_{u_1=0}.$$

  the optimal Lagrange multipliers are the local sensitivies of the optimal value with respect to constraint perturbations.

Primal and dual optimal points

suppose strong duality holds, *i.e.*, we have $p^\star = d^\star$, and both $p^\star$ and $d^\star$ are achieved.

- when can we find a primal optimal point $x^\star$, from a dual optimal point $(\lambda^\star, \nu^\star)$?

consider the Lagrangian, evaluated at $(\lambda^\star, \nu^\star)$,

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

suppose the minimizer of $L(x, \lambda^\star, \nu^\star)$, $\tilde{x}$, is unique. since $x^\star$ minimizes $L(x, \lambda^\star, \nu^\star)$, then we must have $\tilde{x} = x^\star$.

Example: least-norm solution over a polyhedron

consider the problem,

$$\begin{array}{ll} \text{minimize} & x^T x \\ \text{subject to} & Ax \preceq b, \end{array}$$

with $x \in \mathbf{R}^{1000}$, $A \in \mathbf{R}^{10 \times 1000}$. we will assume that A is full rank.
the Lagrangian is

$$L(x, \lambda) = x^T x + \lambda^T (Ax - b),$$

with minimizer $x = -(1/2)A^T \lambda$. the dual problem can be expressed as

$$\begin{array}{ll} \text{maximize} & -(1/4)\lambda^T A A^T \lambda - b^T \lambda \\ \text{subject to} & \lambda \succeq 0. \end{array}$$

- does strong duality hold?
  answer. yes, because we can solve $Ax = b$.

- suppose $\lambda^\star$ is optimal for the dual problem, can we find a primal optimal point? and if so, is it unique?
  answer. yes, $\tilde{x} = -(1/2)A^T\lambda^\star$ minimizes the Lagrangian, but $\tilde{x}$ is unique, so $\tilde{x} = x^\star$.

- for the previous example, we can uniquely construct a primal optimal point $x^\star$, given a dual optimal point $\lambda^\star$.

- how many variables does the primal problem have? **answer.** 1000

- how many variables does the dual problem have? **answer.** 10

- this is one advantage of solving the dual problem and then constructing the solution of the primal problem from the solution of the dual. We will see later however, that if we fully exploit the structure of the primal problem, the time taken to solve these two problems are approximately the same.

Example: linear program

now consider the problem,

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b, \end{array}$$

where A is skinny and full rank. suppose also that strong duality holds, and $p^\star$ and $d^\star$ are achieved. the Lagrangian is,

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

this is bounded if and only if $A^T \lambda + c = 0$. the dual problem can be expressed as,

$$\begin{array}{ll} \text{maximize} & -b^T \lambda \\ \text{subject to} & A^T \lambda + c = 0 \\ & \lambda \succeq 0. \end{array}$$

- can we find a primal optimal point $x^\star$ from a dual optimal point $\lambda^\star$ by minimizing the Lagrangian?
  answer. no, we have,
  $$L(x, \lambda^\star) = -b^T \lambda^\star.$$
  any x minimizes $L(x, \lambda^\star)$, but not every x is optimal for the primal problem.

- can we do this by any other method?
  answer. using complementary slackness, we see that if $\lambda_i^\star > 0$, then we must have $a_i^T x^\star = b_i$. so if the solution to the resulting set of linear equations is unique, it is a primal optimal point.

- it not always possible to construct a primal optimal point from a dual optimal point, even when strong duality holds.

Theorems of alternatives

we want to determine the feasibility of the following system of (not necessarily convex) linear inequalities and equalities

$$f_i(x) \leq 0, \quad i = 1, \ldots, m, \qquad h_i(x) = 0, \quad i = 1, \ldots, p. \tag{1}$$

we will assume the domain $\mathcal{D} = \bigcap_{i=1}^{m} \mathbf{dom}\, f_i \cap \bigcap_{i=1}^{p} \mathbf{dom}\, h_i$ is nonempty. we can write this problem as

$$\begin{array}{ll} \text{minimize} & 0 \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p. \end{array}$$

$$p^\star = \begin{cases} 0 & \text{(1) is feasible} \\ \infty & \text{(1) is infeasible.} \end{cases}$$

the dual function is

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} \left(\sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right)$$

- what happens if there exist $\lambda \succeq 0$, and ν for which $g(\lambda, \nu) > 0$? **answer.** $d^\star = \infty$.

- what happens if there does not exist $\lambda \succeq 0$, and ν for which $g(\lambda, \nu) > 0$? **answer.** $d^\star = 0$.

so,

$$d^\star = \begin{cases} \infty & \lambda \succeq 0,\ g(\lambda,\nu) > 0 \text{ is feasible} \\ 0 & \lambda \succeq 0,\ g(\lambda,\nu) > 0 \text{ is infeasible.} \end{cases}$$

recall $d^\star \leq p^\star$. if (1) is feasible, then $p^\star = 0$, so $d^\star = 0$, which means that the system

$$\lambda \succeq 0, \qquad g(\lambda,\nu) > 0 \tag{2}$$

is infeasible. if (2) is feasible, then $d^\star = \infty$, which means that $p^\star = \infty$, and so (1) is infeasible. note that both systems can be infeasible.

- (1) and (2) are called *weak alternatives*, since at most one of the two is feasible.
- two systems are called *strong alternatives* if exactly one of the two alternatives holds.

- for theorems of alternatives, whether or not the system of inequalities is strict makes a difference. for instance, if we had the system,

$$f_i(x) < 0, \quad i = 1, \ldots, m, \qquad h_i(x) = 0, \quad i = 1, \ldots, p, \tag{3}$$

then (2) would not longer be the weak alternative. instead the weak alternative would be,

$$\lambda \succeq 0, \quad \lambda \neq 0, \quad g(\lambda, \nu) \geq 0. \tag{4}$$

Farkas' lemma

the system of inequalities

$$Ax \preceq 0, \quad c^T x < 0,$$

where $A \in \mathbf{R}^{m \times n}$, $c \in \mathbf{R}^n$, and the system of inequalities

$$A^T y + c = 0, \quad y \succeq 0,$$

are strong alternatives.

- how can we show this?

solution. use LP duality. consider the LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq 0, \end{array}$$

and its dual

$$\begin{array}{ll} \text{maximize} & 0 \\ \text{subject to} & A^T y + c = 0 \\ & y \succeq 0. \end{array}$$

the primal LP is homogeneous, and so the optimal value is 0, if the primal inequality system is not feasible, and $-\infty$ if the the primal inequality system is feasible. the dual has optimal value 0 if the dual inequality system is feasible, and optimal value $-\infty$ if the dual inequality system is not feasible. since $p^\star = d^\star$ (why?), the two systems are strong alternatives.

Application: arbitrage

consider n assets with prices $p_1, p_2, \ldots, p_n$. at the end of the investment period the value of the assets is $v_1, v_2, \ldots, v_n$. $x_1, x_2, \ldots, x_n$ represents the initial investment in each asset ($x_j < 0$ means that we are owe $-x_j$ amount of asset j).

- cost of initial investment is $p^T x$, and final value of the investment is $v^T x$.
- v is uncertain, there are m possible scenarios, $v^{(1)}, \ldots, v^{(m)}$. if scenario i occurs, then the final value of the investment is $v^{(i)T} x$.
- if there is an investment x with $p^T x < 0$, and $v^{(i)T} x \geq 0$, for $i = 1, \ldots, m$, then an *arbitrage* is said to exist.

- in finance, it is often assumed that no arbitrage exists. which means that the system of inequalities

$$Vx \succeq 0, \quad p^T x < 0$$

is infeasible (V is a matrix with rows $v^{(1)T}, \ldots, v^{(m)T}$).

- by Farkas' lemma, the above system is infeasible if and only if there exists y such that

$$V^T y = p, \quad y \succeq 0.$$

- suppose that V is known, and all the prices except the last price p_n is known. we wish to find the set of prices p_n that are consistent with the no-arbitrage assumption. what kind of set is this?
solution. clearly, we must have $p \succeq 0$, which is consistent with intuition. the set must therefore be an interval (it is the projection of a polyhedron onto **R**). we can find this interval by solving a pair of LPs.

to find the minimum p_n we solve,

$$\begin{array}{ll} \text{minimize} & p_n \\ \text{subject to} & V^T y = p \\ & y \succeq 0. \end{array}$$

and to find the maximum p_n, we solve the same LP with maximization.