KTH Mathematics

2E5295/5B5749 Convex optimization with engineering applications

Lecture 4

Linear programming, Lagrangian relaxation and duality

Primal and dual linear programs

For a *primal* linear program

$$(PLP) \qquad \begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b, \\ & x \geq 0, \end{array}$$

we will associate a *dual* linear program

$$(DLP) \qquad \begin{array}{ll} \max & b^T y \\ \text{subject to} & A^T y + s = c, \\ & s \geq 0. \end{array}$$

(We will derive the dual by *Lagrangian relaxation* later.)

Weak duality for linear programming

Proposition. *If x is feasible to (PLP) and y, s is feasible to (DLP), then $c^Tx - b^Ty = x^Ts \geq 0$.*

Proof. Insertion gives the result. $\square$

It follows that *weak duality* holds, i.e., optval$(PLP) \geq$ optval(DLP).

Proposition. *If x is feasible to (PLP) and y, s is feasible to (DLP), and furthermore $x^Ts = 0$, then these solutions are optimal to the primal and the dual, respectively.*

Proof. A consequence of the previous result. $\square$

Note that if $x \geq 0$ and $s \geq 0$, then $x^Ts = 0$ if and only if $x_j s_j = 0$, $j = 1, \ldots, n$.

Strong duality for linear programming

$$
(PLP) \quad \begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^T x \\ \text{subject to} & Ax = b, \\ & x \ge 0. \end{array}
\qquad
(DLP) \quad \begin{array}{ll} \underset{y \in \mathbb{R}^m}{\text{maximize}} & b^T y \\ \text{subject to} & A^T y + s = c, \\ & s \ge 0. \end{array}
$$

Theorem. *If (PLP) has a finite optimal value, then there exist optimal solutions to (PLP) and (DLP). Moreover, the optimal values are equal.*

This result is referred to as *strong duality* for linear programming.

Farkas' lemma

Lemma. *Let A be an $m \times n$ matrix, and let b be an m-dimensional vector. Then exactly one of the following systems has a solution.*

$$1. \quad Ax = b, \; x \geq 0, \qquad\qquad 2. \quad A^T y \geq 0, \; b^T y < 0.$$

Proof. Both systems cannot have a solution simultaneously. (Try.)

It remains to show that if 1 has no solution, then 2 has a solution.

Let $\mathcal{W} = \{w \in \mathbb{R}^m : w = Ax, \; x \geq 0\}$. Then $\mathcal{W}$ is a closed convex set. (Closedness follows by reducing to basic feasible solutions.)

Then 1 has a solution if and only if $b \in \mathcal{W}$.

If $b \notin \mathcal{W}$, then by the separating hyperplane theorem, there is a $y \in \mathbb{R}^m$ such that $y^T b < y^T z$ for all $z \in \mathcal{W}$. This y satisfies 2. $\square$

Proof of strong duality for linear programming

Proof. If (PLP) has optimal value v^*, there is no solution to the system

$$\begin{pmatrix} A & -b \\ -A & b \\ I & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} c^T & -v^* \end{pmatrix}^T \begin{pmatrix} x \\ t \end{pmatrix} < 0.$$

Associating this system with system 2 in Farkas' lemma, simplifications of the corresponding system 1 give the existence of a solution to the system

$$A^T y \leq c, \quad b^T y \geq v^*.$$

Hence, (D) has an optimal solution with value v^*. The existence of a primal optimal solution is proved analogously. □

Linear programming, optimality conditions

Linear program:

$$(PLP) \quad \begin{array}{ll} \min & c^T x \\ \text{subject to} & Ax = b, \\ & x \geq 0. \end{array} \qquad (DLP) \quad \begin{array}{ll} \max & b^T y \\ \text{subject to} & A^T y + s = c, \\ & s \geq 0. \end{array}$$

Optimality conditions:

$$\begin{aligned} Ax &= b, \\ A^T y + s &= c, \\ x_j s_j &= 0, \quad j = 1, \dots, n, \\ x &\geq 0, \\ s &\geq 0. \end{aligned}$$

Relaxation

$$(P) \quad \begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in S. \end{array} \qquad (P_R) \quad \begin{array}{ll} \text{minimize} & f_R(x) \\ \text{subject to} & x \in S_R. \end{array}$$

Definition. *Problem* (P_R) *is a* relaxation *of problem* (P) *if*

$$(i) \;\; S_R \supseteq S, \quad \textit{and} \quad (ii) \;\; f_R(x) \le f(x) \textit{ for all } x \in S.$$

Proposition. *The optimal values satisfy* optval$(P_R) \le$ optval(P).

Proposition. *If* x^* *is a global minimizer to* (P_R) *such that* $x^* \in S$ *and* $f_R(x^*) = f(x^*)$, *then* x^* *is a global minimizer to* (P).

Lagrangian relaxation

$$
(P) \quad \begin{array}{lll}
\text{minimize} & f(x) & \\
\text{subject to} & g_i(x) \geq 0, \quad i \in \mathcal{I}, & \mathcal{I} \bigcup \mathcal{E} = \{1, \dots, m\}, \\
& g_i(x) = 0, \quad i \in \mathcal{E}, & \mathcal{I} \bigcap \mathcal{E} = \emptyset, \\
& x \in X. &
\end{array}
$$

Definition. *For $\lambda \in \mathbb{R}^m$, the* Lagrangian *is $l(x, \lambda) = f(x) - \lambda^T g(x)$.*

For a given $\lambda \in \mathbb{R}^m$ such that $\lambda_i \geq 0$, $i \in \mathcal{I}$, the *Lagrangian relaxation problem* is given by

$$
(P_\lambda) \quad \begin{array}{ll}
\text{minimize} & f(x) - \lambda^T g(x) \\
\text{subject to} & x \in X.
\end{array}
$$

NB! λ is fixed when solving (P_λ), x is the variable.

Lagrangian duality

The *Lagrangian dual* problem (D) is obtained by making the Lagrangian relaxation as strong as possible.

$$(D) \quad \begin{array}{ll} \text{maximize} & \varphi(\lambda) \\ \text{subject to} & \lambda \in \mathbb{R}^m, \quad \lambda_i \geq 0, i \in \mathcal{I}, \end{array}$$

where

$$\varphi(\lambda) = \underset{x \in X}{\text{minimize}}\ f(x) - \lambda^T g(x).$$

Weak duality. *The optimal values satisfy* optval$(D) \leq$ optval(P). *The difference* optval$(P) -$ optval(D) *is referred to as the* duality gap.

The dual of a linear program

$$(PLP) \qquad \begin{array}{ll} \underset{x \in I\!R^n}{\text{minimize}} & c^T x \\ \text{subject to} & Ax = b, \\ & x \geq 0. \end{array}$$

For a given $y \in I\!R^m$, we obtain

$$\varphi(y) = \underset{x \geq 0}{\text{minimize}}\ c^T x - y^T(Ax - b) = \begin{cases} b^T y & \text{if } A^T y \leq c, \\ -\infty & \text{otherwise.} \end{cases}$$

Consequently,

$$(DLP) \qquad \begin{array}{ll} \underset{y \in I\!R^m}{\text{maximize}} & b^T y \\ \text{subject to} & A^T y \leq c. \end{array}$$

Example of linear program: Curve fitting

Assume that a line $y = kx + l$ is to be fit to a set of given points (x_i, y_i), $i = 1, \dots, m$.

Consider to ways of fitting the line:

- Choose k and l so that the maximum deviation in the y-direction is minimized, i.e., let k and l solve $\min_{k,l}\{\max_i |kx_i + l - y_i|\}$.
- Choose k and l so that the sum of the deviations in the y-direction is minimized, i.e., let k and l solve $\min_{k,l} \sum_i |kx_i + l - y_i|$.

Both these curve fitting problems can be formulated as linear programs.

Suggested reading

Suggested reading in the textbook:

- Sections 4.1–4.3.
- Sections 5.1–5.2.

Linearly constrained convex program

We will consider a linearly constrained convex optimization problem on the form

$$(LCP) \quad \begin{array}{lll} \underset{x \in I\!R^n}{\text{minimize}} & f(x) & \\ \text{subject to} & a_i^T x - b_i \geq 0, & i \in \mathcal{I}, \\ & a_i^T x - b_i = 0, & i \in \mathcal{E}, \end{array} \qquad \begin{array}{l} \mathcal{I} \bigcup \mathcal{E} = \{1, \dots, m\} \\ \mathcal{I} \bigcap \mathcal{E} = \emptyset, \end{array}$$

where $f : I\!R^n \to I\!R$ is convex and twice continuously differentiable.

Let $F = \{x \in I\!R^n : a_i^T x - b_i \geq 0, \ i \in \mathcal{I}, a_i^T x - b_i = 0, \ i \in \mathcal{E}\}$.

Definition. *A direction p is a* feasible direction *to F in x^* if there is $\bar{\alpha} > 0$ such that $x^* + \alpha p \in F$ for $\alpha \in [0, \bar{\alpha}]$.*

Definition. *A direction p is a* descent direction *to f in x^* if $\nabla f(x^*)^T p < 0$.*

Optimality conditions for linearly constrained convex program

Let $\mathcal{A}(x^*)$ denote the set of constraints that are active at $x^* \in F$, i.e., $\mathcal{A}(x^*) = \{i \in \{1, \ldots, n\}$ such that $a_i^T x^* - b_i = 0\}$.

Then, $x^* \in F$ is a minimizer to (LCP) if and only if there is no feasible descent direction.

Proposition. *The point $x^* \in F$ is a minimizer to (LCP) if and only if it is a minimizer to*

$$(LP) \qquad \begin{array}{ll} \underset{p \in \mathbb{R}^n}{\text{minimize}} & \nabla f(x^*)^T p \\ \text{subject to} & a_i^T p \geq 0, \quad i \in \mathcal{I} \cap \mathcal{A}(x^*), \\ & a_i^T p = 0, \quad i \in \mathcal{E}. \end{array}$$

Optimality conditions for linearly constrained convex program, cont.

Hence, if x^* is a minimizer, the linear programs

$$\begin{array}{lll} \underset{p \in \mathbb{R}^n}{\text{minimize}} & \nabla f(x^*)^T p & \\ \text{subject to} & a_i^T p \geq 0, & i \in \mathcal{I} \cap \mathcal{A}(x^*), \\ & a_i^T p = 0, & i \in \mathcal{E}, \end{array}$$

and

$$\begin{array}{ll} \text{maximize} & \sum_{i \in \mathcal{A}(x^*)} 0 \lambda_i \\ \text{subject to} & \sum_{i \in \mathcal{A}(x^*)} a_i \lambda_i = \nabla f(x^*), \\ & \lambda_i \geq 0, \ i \in \mathcal{I} \cap \mathcal{A}(x^*). \end{array}$$

(The second problem is the LP-dual of the first one.)

Optimality conditions for linearly constrained convex program, cont.

The first-order necessary and sufficient optimality conditions for the convex linearly constrained optimization problem can now be stated.

Proposition. *The point* x^* *is a minimizer to* (LCP) *if and only if there is a* $\lambda^* \in \mathbb{R}^m$ *such that*

(i) $a_i^T x^* - b_i \geq 0$, $i \in \mathcal{I}$, *and* $a_i^T x^* - b_i = 0$, $i \in \mathcal{E}$,

(ii) $\nabla f(x^*) = \sum_{i=1}^m a_i \lambda_i^* = A^T \lambda^*$,

(iii) $\lambda_i^* \geq 0$, $i \in \mathcal{I}$,

(iv) $\lambda_i^*(a_i^T x^* - b_i) = 0, \ i \in \mathcal{I}$.

These conditions are often referred to as the *KKT conditions*.

Optimality conditions for linearly constrained convex program, cont.

Let $l(x,\lambda) = f(x) - \lambda^T(Ax - b)$, let $\varphi(\lambda) = \min_{x \in I\!R^n} l(x,\lambda)$, and let

$$(D) \quad \begin{array}{ll} \text{maximize} & \varphi(\lambda) \\ \text{subject to} & \lambda \in I\!R^m, \quad \lambda_i \geq 0, i \in \mathcal{I}. \end{array}$$

(i) states that x^* is feasible to (LCP).

(ii) states that x^* solves $\min_{x \in I\!R^n} l(x,\lambda^*)$.

(iii) states that λ^* is feasible to (D).

(iv) states that $f(x^*) = l(x^*,\lambda^*) = \varphi(\lambda^*)$.

Consequently, the first-order optimality conditions are equivalent to strong duality.