Lecture 7
Duality

- Lagrange dual function
- Lagrange dual problem
- Strong duality and Slater’s condition
- KKT optimality conditions
- Sensitivity analysis
- Equality constraints
- Generalized inequalities
- Theorems of alternatives
standard form problem (without equality constraints)

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_i(x) \leq 0, \ i = 1, \ldots, m
\end{align*}
\]

• optimal value \( p^* \), domain \( D \)

• called **primal problem** (in context of duality)

(for now) we **don’t** assume convexity

**Lagrangian** \( L : \mathbb{R}^{n+m} \rightarrow \mathbb{R} \)

\[
L(x, \lambda) = f_0(x) + \lambda_1 f_1(x) + \cdots + \lambda_m f_m(x)
\]

• \( \lambda_i \) called **Lagrange multipliers** or **dual variables**

• objective is **augmented** with weighted sum of constraint functions
Lagrange dual function

(Lagrange) dual function $g : \mathbb{R}^m \rightarrow \mathbb{R} \cup \{-\infty\}$

\[
g(\lambda) = \inf_x L(x, \lambda) = \inf_x (f_0(x) + \lambda_1 f_1(x) + \cdots + \lambda_m f_m(x))
\]

- minimum of augmented cost as function of weights
- can be $-\infty$ for some $\lambda$
- $g$ is concave (even if $f_i$ not convex!)

example: LP

\[
\begin{align*}
\text{minimize} & & c^T x \\
\text{subject to} & & a_i^T x - b_i \leq 0, \ i = 1, \ldots, m
\end{align*}
\]

\[
L(x, \lambda) = c^T x + \sum_{i=1}^{m} \lambda_i (a_i^T x - b_i) = -b^T \lambda + (A^T \lambda + c)^T x
\]

hence $g(\lambda) = \begin{cases} 
- b^T \lambda & \text{if } A^T \lambda + c = 0 \\
- \infty & \text{otherwise}
\end{cases}$
**Lower bound property**

if $\lambda \geq 0$ and $x$ is primal feasible, then

$$g(\lambda) \leq f_0(x)$$

**proof:** if $f_i(x) \leq 0$ and $\lambda_i \geq 0$,

$$f_0(x) \geq f_0(x) + \sum_i \lambda_i f_i(x)$$

$$\geq \inf_z \left( f_0(z) + \sum_i \lambda_i f_i(z) \right)$$

$$= g(\lambda)$$

$f_0(x) - g(\lambda)$ is called the **duality gap** of (primal feasible) $x$ and $\lambda \geq 0$

minimize over primal feasible $x$ to get, for any $\lambda \geq 0$,

$$g(\lambda) \leq p^*$$

$\lambda \in \mathbb{R}^m$ is **dual feasible** if $\lambda \geq 0$ and $g(\lambda) > -\infty$

dual feasible points yield lower bounds on optimal value!
Let’s find best lower bound on $p^*$:

\[
\begin{align*}
\text{maximize} & \quad g(\lambda) \\
\text{subject to} & \quad \lambda \geq 0
\end{align*}
\]

- called \textbf{(Lagrange) dual problem} (associated with primal problem)
- always a convex problem, even if primal isn’t!
- optimal value denoted $d^*$
- we always have $d^* \leq p^*$ (called \textit{weak duality})
- $p^* - d^*$ is \textit{optimal duality gap}
Strong duality

for convex problems, we (usually) have strong duality:

\[ d^* = p^* \]

when strong duality holds, dual optimal \( \lambda^* \) serves as certificate of optimality for primal optimal point \( x^* \)

many conditions or constraint qualifications guarantee strong duality for convex problems

Slater’s condition: if primal problem is strictly feasible (and convex), i.e., there exists \( x \in \text{relin}t \ D \) with

\[ f_i(x) < 0, \ i = 1, \ldots, m \]

then we have \( p^* = d^* \)
Dual of linear program

(primal) LP

minimize $c^T x$
subject to $Ax \leq b$

- $n$ variables, $m$ inequality constraints

dual of LP is (after making implicit equality constraints explicit)

maximize $-b^T \lambda$
subject to $A^T \lambda + c = 0$
$\lambda \geq 0$

- dual of LP is also an LP (indeed, in std LP format)

- $m$ variables, $n$ equality constraints, $m$ nonnegativity constraints

for LP we have strong duality except in one (pathological) case: primal and dual both infeasible ($p^* = +\infty$, $d^* = -\infty$)
Dual of quadratic program

(primal) QP

\[
\begin{align*}
\text{minimize} & \quad x^T P x \\
\text{subject to} & \quad Ax \leq b
\end{align*}
\]

we assume \( P \succ 0 \) for simplicity

Lagrangian is \( L(x, \lambda) = x^T P x + \lambda^T (Ax - b) \)

\( \nabla_x L(x, \lambda) = 0 \) yields \( x = -(1/2) P^{-1} A^T \lambda \), hence dual function is

\[
g(\lambda) = -(1/4) \lambda^T A P^{-1} A^T \lambda - b^T \lambda
\]

• concave quadratic function

• all \( \lambda \geq 0 \) are dual feasible

dual of QP is

\[
\begin{align*}
\text{maximize} & \quad -(1/4) \lambda^T A P^{-1} A^T \lambda - b^T \lambda \\
\text{subject to} & \quad \lambda \geq 0
\end{align*}
\]

... another QP

Duality
Duality in algorithms

many algorithms produce at iteration $k$

- a primal feasible $x^{(k)}$

- and a dual feasible $\lambda^{(k)}$

with $f_0(x^{(k)}) - g(\lambda^{(k)}) \to 0$ as $k \to \infty$

hence at iteration $k$ we know $p^* \in [g(\lambda^{(k)}), f_0(x^{(k)})]$ 

- useful for stopping criteria 

- algorithms that use dual solution are often more efficient (e.g., LP)
Nonheuristic stopping criteria

**absolute error**

\[
\text{absolute error} = f_0(x^{(k)}) - p^* \leq \epsilon
\]

stopping criterion:

\[
\text{until } \left( f_0(x^{(k)}) - g(\lambda^{(k)}) \leq \epsilon \right)
\]

**relative error**

\[
\text{relative error} = \frac{f_0(x^{(k)}) - p^*}{|p^*|} \leq \epsilon
\]

stopping criterion:

\[
\text{until } \left( g(\lambda^{(k)}) > 0 \quad \& \quad \frac{f_0(x^{(k)}) - g(\lambda^{(k)})}{g(\lambda^{(k)})} \leq \epsilon \right)
\]

\[
\text{or } \left( f_0(x^{(k)}) < 0 \quad \& \quad \frac{f_0(x^{(k)}) - g(\lambda^{(k)})}{-f_0(x^{(k)})} \leq \epsilon \right)
\]

achieve **target value** \( \ell \) or, prove \( \ell \) is unachievable (i.e., determine either \( p^* \leq \ell \) or \( p^* > \ell \))

stopping criterion:

\[
\text{until } \left( f_0(x^{(k)}) \leq \ell \text{ or } g(\lambda^{(k)}) > \ell \right)
\]
Complementary slackness

suppose $x^*$, $\lambda^*$ are primal, dual feasible with zero duality gap (hence, they are primal, dual optimal)

$$f_0(x^*) = g(\lambda^*)$$

$$= \inf_x \left( f_0(x) + \sum_{i=1}^{m} \lambda^*_i f_i(x) \right)$$

$$\leq f_0(x^*) + \sum_{i=1}^{m} \lambda^*_i f_i(x^*)$$

hence we have $\sum_{i=1}^{m} \lambda^*_i f_i(x^*) = 0$, and so

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

• called complementary slackness condition

• $i$th constraint inactive at optimum $\implies \lambda_i = 0$

• $\lambda^*_i > 0$ at optimum $\implies$ $i$th constraint active at optimum
suppose

- $f_i$ are differentiable

- $x^*, \lambda^*$ are (primal, dual) optimal, with zero duality gap

by complementary slackness we have

$$f_0(x^*) + \sum_i \lambda_i^* f_i(x^*) = \inf_x \left( f_0(x) + \sum_i \lambda_i^* f_i(x) \right)$$

i.e., $x^*$ minimizes $L(x, \lambda^*)$

therefore

$$\nabla f_0(x^*) + \sum_i \lambda_i^* \nabla f_i(x^*) = 0$$
so if \( x^*, \lambda^* \) are (primal, dual) optimal, with zero duality gap, they satisfy

\[
\begin{align*}
  f_i(x^*) & \leq 0 \\
  \lambda_i^* & \geq 0 \\
  \lambda_i^* f_i(x^*) & = 0 \\
  \nabla f_0(x^*) + \sum_i \lambda_i^* \nabla f_i(x^*) & = 0
\end{align*}
\]

the **Karush-Kuhn-Tucker** (KKT) optimality conditions

conversely, if the problem is convex and \( x^*, \lambda^* \) satisfy KKT, then they are (primal, dual) optimal
Geometric interpretation of duality

consider set

\[ A = \{ (u, t) \in \mathbb{R}^{m+1} | \exists x \ f_i(x) \leq u_i, \ f_0(x) \leq t \} \]

- \( A \) is convex if \( f_i \) are
- for \( \lambda \geq 0, \)

\[
g(\lambda) = \inf \left\{ \left[ \begin{array}{c} \lambda \\ 1 \end{array} \right]^T \left[ \begin{array}{c} u \\ t \end{array} \right] \bigg| \left[ \begin{array}{c} u \\ t \end{array} \right] \in A \right\}
\]
(Idea of) proof of Slater’s theorem

problem convex, strictly feasible $\implies$ strong duality

\[ (0, p^*) \in \partial \mathcal{A} \implies \exists \text{ supporting hyperplane at } (0, p^*): \]
\[ (u, t) \in \mathcal{A} \implies \mu_0 (t - p^*) + \mu^T u \geq 0 \]

- $\mu_0 \geq 0$, $\mu \geq 0$, $(\mu, \mu_0) \neq 0$
- strong duality $\iff \exists \text{ supp. hyperpl. with } \mu_0 > 0$:
  for $\lambda^* = \mu / \mu_0$, we have
  \[ p^* \leq t + \lambda^T u \ orall (t, u) \in \mathcal{A} \]
  \[ p^* \leq g(\lambda^*) \]

- Slater’s condition: there exists $(u, t) \in \mathcal{A}$ with $u \prec 0$; implies that all supporting hyperplanes at $(0, p^*)$ are non-vertical ($\mu_0 > 0$)
Sensitivity analysis via duality

define $p^*(u)$ as the optimal value of

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_i(x) \leq u_i, \quad i = 1, \ldots, m
\end{align*}
\]

\(\lambda^*\) gives lower bound on $p^*(u)$:

\[
p^*(u) \geq p^* - \sum_{i=1}^{m} \lambda^*_i u_i
\]

- if $\lambda^*_i$ large: $u_i < 0$ greatly increases $p^*$
- if $\lambda^*_i$ small: $u_i > 0$ does not decrease $p^*$ too much

if $p^*(u)$ is differentiable, $\lambda^*_i = -\frac{\partial p^*(0)}{\partial u_i}$

$\lambda^*_i$ is sensitivity of $p^*$ w.r.t. $i$th constraint
Equality constraints

\[ \text{minimize } \quad f_0(x) \]

subject to \[ f_i(x) \leq 0, \quad i = 1, \ldots, m \]
\[ h_i(x) = 0, \quad i = 1, \ldots, p \]

• optimal value \( p^* \)

• again assume (for now) not necessarily convex

define \textbf{Lagrangian} \( L : \mathbb{R}^{n+m+p} \rightarrow \mathbb{R} \) as

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

\textbf{dual function} is \( g(\lambda, \nu) = \inf_x L(x, \lambda, \nu) \)

\((\lambda, \nu)\) is dual feasible if \( \lambda \succeq 0 \) and \( g(\lambda, \nu) > -\infty \)

(no sign condition on \( \nu \))
**lower bound property:** if $x$ is primal feasible and $(\lambda, \nu)$ is dual feasible, then $g(\lambda, \nu) \leq f_0(x)$, hence

$$g(\lambda, \nu) \leq p^*$$

**dual problem:** find best lower bound

maximize $g(\lambda, \nu)$

subject to $\lambda \geq 0$

(note $\nu$ unconstrained), optimal value $d^*$

**weak duality:** $d^* \leq p^*$ always

**strong duality:** if primal is convex then (usually) $d^* = p^*$

**Slater condition:** if primal is convex (i.e., $f_i$ convex, $h_i$ affine) and strictly feasible, i.e., there exists $x \in \text{relint } D$ s.t.

$$f_i(x) < 0, \quad h_i(x) = 0,$$

then $d^* = p^*$
Example: equality constrained least-squares

minimize $x^T x$
subject to $Ax = b$

$A$ is fat, full rank (solution is $x^* = A^T( AA^T)^{-1} b$)

dual function is

$$g(\nu) = \inf \left( x^T x + \nu^T (Ax - b) \right) = -\frac{1}{4} \nu^T AA^T \nu - b^T \nu$$

dual problem is

maximize $-\frac{1}{4} \nu^T AA^T \nu - b^T \nu$

solution: $\nu^* = -2( AA^T )^{-1} b$

can check $d^* = p^*$
KKT optimality conditions

assume \( f_i, h_i \) differentiable

if \( x^*, \lambda^*, \nu^* \) are optimal, with zero duality gap, then they satisfy KKT conditions

\[
\begin{align*}
& f_i(x^*) \leq 0, \quad h_i(x^*) = 0 \\
& \lambda_i^* \geq 0 \\
& \lambda_i^* f_i(x^*) = 0 \\
& \nabla f_0(x^*) + \sum_i \lambda_i^* \nabla f_i(x^*) + \sum_i \nu_i^* \nabla h_i(x^*) = 0
\end{align*}
\]

conversely, if they satisfy KKT and the problem is convex, then \( x^*, \lambda^*, \nu^* \) are optimal

example:
optimality conditions for equality constraints only

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad Ax = b
\end{align*}
\]

\( x^* \) optimal \( \iff \) \( \exists \nu^* \) s.t. \( \nabla f_0(x^*) + A^T \nu^* = 0 \)
Introducing equality constraints

**idea:** simple transformation of primal problem can lead to very different dual

**example:** unconstrained geometric programming

**primal problem:**

$$\text{minimize } \log \sum_{i=1}^{m} \exp(a_i^T x - b_i)$$

dual function is constant $g = p^*$

(we have strong duality, but it’s useless)

now **rewrite primal problem** as

$$\text{minimize } \log \sum_{i=1}^{m} \exp y_i$$

subject to $y = Ax - b$

- introduce $m$ new vbles $y_1, \ldots, y_m$

- introduce $m$ new equality constraints $y = Ax - b$
dual function

\[ g(\nu) = \inf_{x,y} \left( \log \sum_{i=1}^{m} \exp y_i + \nu^T (Ax - b - y) \right) \]

- infimum is \(-\infty\) if \(A^T \nu \neq 0\)

- assuming \(A^T \nu = 0\), let’s minimize over \(y\):

\[
\frac{e^{y_i}}{\sum_{j=1}^{n} e^{y_j}} = \nu_i
\]

solvable iff \(\nu_i > 0\), \(1^T \nu = 1\)

\[ g(\nu) = -\sum_{i} \nu_i \log \nu_i - b^T \nu \]

- same expression if \(\nu \geq 0\), \(1^T \nu = 1\) \((0 \log 0 = 0)\)

dual problem

maximize \(-b^T \nu - \sum_{i} \nu_i \log \nu_i\)

subject to \(1^T \nu = 1\), \((\nu \geq 0)\)
\(A^T \nu = 0\)

moral: trivial reformulation can yield different dual
Generalized inequalities

minimize \( f_0(x) \)
subject to \( f_i(x) \leq K_i \ 0, \ i = 1, \ldots, L \)

- \( \leq K_i \) are generalized inequalities on \( \mathbb{R}^{m_i} \)

- \( f_i : \mathbb{R}^n \rightarrow \mathbb{R}^{m_i} \) are \( K_i \)-convex

Lagrangian \( L : \mathbb{R}^n \times \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_L} \rightarrow \mathbb{R} \),

\[
L(x, \lambda_1, \ldots, \lambda_L) = f_0(x) + \lambda_1^T f_1(x) + \cdots + \lambda_L^T f_L(x)
\]

dual function

\[
g(\lambda_1, \ldots, \lambda_L) = \inf_x (f_0(x) + \lambda_1^T f_1(x) + \cdots + \lambda_L^T f_L(x))
\]

\( \lambda_i \) dual feasible if \( \lambda_i \leq K_i^* 0, \ g(\lambda_1, \ldots, \lambda_L) > -\infty \)

lower bound property: if \( x \) primal feasible and \( (\lambda_1, \ldots, \lambda_L) \) is dual feasible, then

\[
g(\lambda_1, \ldots, \lambda_L) \leq f_0(x)
\]

(hence, \( g(\lambda_1, \ldots, \lambda_L) \leq p^* \))
dual problem

maximize \[ g(\lambda_1, \ldots, \lambda_L) \]
subject to \[ \lambda_i \geq K_i^* \quad 0, \quad i = 1, \ldots, L \]

weak duality: \( d^* \leq p^* \) always

strong duality: \( d^* = p^* \) usually

Slater condition: if primal is strictly feasible, i.e.,

\[ \exists x \in \text{relint } D : f_i(x) \prec K_i 0, \quad i = 1, \ldots, L \]

then \( d^* = p^* \)
Example: semidefinite programming

minimize \( c^T x \)
subject to \( F_0 + x_1 F_1 + \cdots + x_n F_n \preceq 0 \)

**Lagrangian** (multiplier \( Z = Z^T \in \mathbb{R}^{m \times m} \))

\[
L(x, Z) = c^T x + \text{Tr} \ Z(F_0 + x_1 F_1 + \cdots + x_n F_n)
\]

dual function

\[
g(Z) = \inf_x \left( c^T x + \text{Tr} \ Z(F_0 + x_1 F_1 + \cdots + x_n F_n) \right)
\]

\[
= \begin{cases} 
\text{Tr} \ F_0 Z & \text{if } \text{Tr} \ F_i Z + c_i = 0, \ i = 1, \ldots, n \\
-\infty & \text{otherwise}
\end{cases}
\]

dual problem

maximize \( \text{Tr} \ F_0 Z \)
subject to \( \text{Tr} \ F_i Z + c_i = 0, \ i = 1, \ldots, n \)
\( Z = Z^T \succeq 0 \)

**strong duality** holds if there exists \( x \) with

\[
F_0 + x_1 F_1 + \cdots + x_n F_n \prec 0
\]
Theorem of alternatives

\( f_1, \ldots, f_m \) convex with \( \text{dom} f_i = \mathbb{R}^n \)

exactly one of the following is true:

1. there exists \( x \) with \( f_i(x) < 0 \), \( i = 1, \ldots, m \)

2. there exists \( \lambda \neq 0 \) with \( \lambda \geq 0 \),

\[
g(\lambda) = \inf_x (\lambda_1 f_1(x) + \cdots + \lambda_m f_m(x)) \geq 0
\]

- called **alternatives**

- use in practice: \( \lambda \) that satisfies 2nd condition proves \( f_i(x) < 0 \) is infeasible

example: \( f_i(x) = a_i^T x - b_i \)

1. there exists \( x \) with \( Ax < b \)

2. there exists \( \lambda \geq 0, \lambda \neq 0, b^T \lambda \leq 0, A^T \lambda = 0 \)
**proof.** From convex duality:

**primal problem**

\[
\begin{align*}
& \text{minimize} & & t \\
& \text{subject to} & & f_i(x) \leq t, \ i = 1, \ldots, m
\end{align*}
\]

(variables \(x, t\))

**dual problem**

\[
\begin{align*}
& \text{maximize} & & g(\lambda) \\
& \text{subject to} & & \lambda \succeq 0 \\
& & & 1^T \lambda = 1
\end{align*}
\]

- Slater’s condition is satisfied, hence \(p^* = d^*\)
- 1st alternative: \(\iff p^* < 0\)
- 2nd alternative: \(\iff p^* \geq 0\)