5. Duality

- Lagrange dual problem
- weak and strong duality
- geometric interpretation
- optimality conditions
- perturbation and sensitivity analysis
- examples
- generalized inequalities
Lagrangian

**standard form problem** (not necessarily convex)

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

variable \( x \in \mathbb{R}^n \), domain \( D \), optimal value \( p^* \)

**Lagrangian:** \( L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R} \), with \( \text{dom} L = D \times \mathbb{R}^m \times \mathbb{R}^p \),

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

- weighted sum of objective and constraint functions
- \( \lambda_i \) is Lagrange multiplier associated with \( f_i(x) \leq 0 \)
- \( \nu_i \) is Lagrange multiplier associated with \( h_i(x) = 0 \)
Lagrange dual function

Lagrange dual function: \( g : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R} \),

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

\( g \) is concave, can be \(-\infty\) for some \( \lambda, \nu \)

**lower bound property:** if \( \lambda \succeq 0 \), then \( g(\lambda, \nu) \leq p^* \)

proof: if \( \tilde{x} \) is feasible and \( \lambda \succeq 0 \), then

\[
f_0(\tilde{x}) \geq L(\tilde{x}, \lambda, \nu) \geq \inf_{x \in D} L(x, \lambda, \nu) = g(\lambda, \nu)
\]

minimizing over all feasible \( \tilde{x} \) gives \( p^* \geq g(\lambda, \nu) \)
Least-norm solution of linear equations

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

dual function

- Lagrangian is \( L(x, \nu) = x^T x + \nu^T (Ax - b) \)
- to minimize \( L \) over \( x \), set gradient equal to zero:

\[
\nabla_x L(x, \nu) = 2x + A^T \nu = 0 \quad \Rightarrow \quad x = -(1/2) A^T \nu
\]

- plug in in \( L \) to obtain \( g \):

\[
g(\nu) = L((-1/2) A^T \nu, \nu) = -\frac{1}{4} \nu^T A A^T \nu - b^T \nu
\]

a concave function of \( \nu \)

lower bound property: \( p^* \geq -(1/4) \nu^T A A^T \nu - b^T \nu \) for all \( \nu \)
Standard form LP

\[
\begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to} & \quad Ax = b, \quad x \geq 0
\end{align*}
\]

dual function

- Lagrangian is

\[
L(x, \lambda, \nu) = c^T x + \nu^T (Ax - b) - \lambda^T x
\]

\[
= -b^T \nu + (c + A^T \nu - \lambda)^T x
\]

- \( L \) is affine in \( x \), hence

\[
g(\lambda, \nu) = \inf_x L(x, \lambda, \nu) = \left\{ \begin{array}{ll} 
- b^T \nu & A^T \nu - \lambda + c = 0 \\
-\infty & \text{otherwise}
\end{array} \right.
\]

\( g \) is linear on affine domain \( \{(\lambda, \nu) \mid A^T \nu - \lambda + c = 0\} \), hence concave

**lower bound property:** \( p^* \geq -b^T \nu \) if \( A^T \nu + c \geq 0 \)
Equality constrained norm minimization

minimize \|x\|
subject to \ Ax = b

dual function

\[ g(\nu) = \inf_x (\|x\| - \nu^T Ax + b^T \nu) = \begin{cases} 
  b^T \nu & \|A^T \nu\|_* \leq 1 \\
  -\infty & \text{otherwise}
\end{cases} \]

where \( \|\nu\|_* = \sup\|u\|\leq 1 u^T \nu \) is dual norm of \( \| \cdot \| \)

proof: follows from \( \inf_x (\|x\| - y^T x) = 0 \) if \( \|y\|_* \leq 1 \), \(-\infty \) otherwise

\begin{itemize}
  \item if \( \|y\|_* \leq 1 \), then \( \|x\| - y^T x \geq 0 \) for all \( x \), with equality if \( x = 0 \)
  \item if \( \|y\|_* > 1 \), choose \( x = tu \) where \( \|u\| \leq 1 \), \( u^T y = \|y\|_* > 1 \):
    \[ \|x\| - y^T x = t(\|u\| - \|y\|_*) \rightarrow -\infty \text{ as } t \rightarrow \infty \]
\end{itemize}

lower bound property: \( p^* \geq b^T \nu \) if \( \|A^T \nu\|_* \leq 1 \)
Two-way partitioning

minimize \( x^T W x \)
subject to \( x_i^2 = 1, \quad i = 1, \ldots, n \)

- a nonconvex problem; feasible set contains \( 2^n \) discrete points
- interpretation: partition \( \{1, \ldots, n\} \) in two sets; \( W_{ij} \) is cost of assigning \( i, j \) to the same set; \(-W_{ij}\) is cost of assigning to different sets

**dual function**

\[
g(\nu) = \inf_x (x^T W x + \sum_i \nu_i (x_i^2 - 1)) = \inf_x x^T (W + \text{diag}(\nu)) x - 1^T \nu
\]

\[
= \begin{cases} 
-1^T \nu & W + \text{diag}(\nu) \succeq 0 \\
-\infty & \text{otherwise}
\end{cases}
\]

**lower bound property:** \( p^* \geq -1^T \nu \) if \( W + \text{diag}(\nu) \succeq 0 \)

example: \( \nu = -\lambda_{\min}(W) 1 \) gives bound \( p^* \geq n \lambda_{\min}(W) \)
Lagrange dual and conjugate function

minimize \( f_0(x) \)
subject to \( Ax \leq b, \ Cx = d \)

dual function

\[
g(\lambda, \nu) = \inf_{x \in \text{dom} f_0} \left( f_0(x) + (A^T \lambda + C^T \nu)^T x - b^T \lambda - d^T \nu \right)
\]

\[
= -f_0^*(-A^T \lambda - C^T \nu) - b^T \lambda - d^T \nu
\]

- recall definition of conjugate \( f^*(y) = \sup_{x \in \text{dom} f} (y^T x - f(x)) \)
- simplifies derivation of dual if conjugate of \( f_0 \) is known

example: entropy maximization

\[
f_0(x) = \sum_{i=1}^{n} x_i \log x_i, \quad f_0^*(y) = \sum_{i=1}^{n} e^{y_i - 1}
\]
The dual problem

Lagrange dual problem

maximize \( g(\lambda, \nu) \)
subject to \( \lambda \succeq 0 \)

• finds best lower bound on \( p^* \), obtained from Lagrange dual function
• a convex optimization problem; optimal value denoted \( d^* \)
• \( \lambda, \nu \) are dual feasible if \( \lambda \succeq 0, (\lambda, \nu) \in \text{dom} \ g \)
• often simplified by making implicit constraint \( (\lambda, \nu) \in \text{dom} \ g \) explicit

example: standard form LP and its dual (page 1–5)

minimize \( c^T x \)
subject to \( Ax = b \)
\( x \succeq 0 \)

maximize \( -b^T \nu \)
subject to \( A^T \nu + c \succeq 0 \)
Weak and strong duality

**Weak duality:** \( d^* \leq p^* \)

- always holds (for convex and nonconvex problems)
- can be used to find nontrivial lower bounds for difficult problems
  
  for example, solving the SDP

\[
\begin{align*}
\text{maximize} & \quad -1^T \nu \\
\text{subject to} & \quad W + \text{diag}(\nu) \succeq 0
\end{align*}
\]

  gives a lower bound for the two-way partitioning problem on page 1–7

**Strong duality:** \( d^* = p^* \)

- does not hold in general
- (usually) holds for convex problems
- conditions that guarantee strong duality in convex problems are called **constraint qualifications**
Slater’s constraint qualification

strong duality holds for a convex problem

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

if it is strictly feasible, \textit{i.e.},

\[
\exists x \in \text{int} \, \mathcal{D} : \quad f_i(x) < 0, \quad i = 1, \ldots, m, \quad Ax = b
\]

\begin{itemize}
  \item also guarantees that the dual optimum is attained (if \( p^* > -\infty \))
  \item can be sharpened: \textit{e.g.,} can replace \text{int} \, \mathcal{D} with \text{relint} \, \mathcal{D} (interior relative to affine hull); linear inequalities do not need to hold with strict inequality, \ldots
  \item there exist many other types of constraint qualifications
\end{itemize}
Inequality form LP

primal problem

minimize \( c^T x \)
subject to \( Ax \leq b \)

dual function

\[
g(\lambda) = \inf_x \left( (c + A^T \lambda)^T x - b^T \lambda \right) = \begin{cases} -b^T \lambda & A^T \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}
\]

dual problem

maximize \(-b^T \lambda\)
subject to \(A^T \lambda + c = 0, \quad \lambda \succeq 0\)

• from Slater’s condition: \( p^* = d^* \) if \( A\tilde{x} \prec b \) for some \( \tilde{x} \)

• in fact, \( p^* = d^* \) except when primal and dual are infeasible
Quadratic program

primal problem (assume $P \in \mathbb{S}^{++}_n$)

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

dual function

\[
g(\lambda) = \inf_x \left( x^T P x + \lambda^T (Ax - b) \right) = -\frac{1}{4} \lambda^T A P^{-1} A^T \lambda - b^T \lambda
\]

dual problem

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

- from Slater’s condition: $p^* = d^*$ if $A\tilde{x} \prec b$ for some $\tilde{x}$
- in fact, $p^* = d^*$ always
A nonconvex problem with strong duality

minimize \( x^T A x + 2b^T x \)
subject to \( x^T x \leq 1 \)

\( A \not\succeq 0 \), hence nonconvex

**dual function:** \( g(\lambda) = \inf_x (x^T (A + \lambda I)x + 2b^T x - \lambda) \)

- unbounded below if \( A + \lambda I \not\succeq 0 \) or if \( A + \lambda I \succeq 0 \) and \( b \not\in \mathcal{R}(A + \lambda I) \)
- minimized by \( x = -(A + \lambda I)^\dagger b \) otherwise: \( g(\lambda) = -b^T (A + \lambda I)^\dagger b - \lambda \)

**dual problem** and equivalent SDP:

maximize \( -b^T (A + \lambda I)^\dagger b - \lambda \)
subject to
- \( A + \lambda I \succeq 0 \)
- \( b \in \mathcal{R}(A + \lambda I) \)

maximize \( -t - \lambda \)
subject to \( \begin{bmatrix} A + \lambda I & b \\ b^T & t \end{bmatrix} \succeq 0 \)

strong duality although primal problem is not convex (not easy to show)
Geometric interpretation

for simplicity, consider problem with one constraint $f_1(x) \leq 0$

interpretation of dual function:

$$g(\lambda) = \inf_{(u,t) \in G} (t + \lambda u), \quad \text{where} \quad G = \{(f_1(x), f_0(x)) \mid x \in D\}$$

- $\lambda u + t = g(\lambda)$ is (non-vertical) supporting hyperplane to $G$
- hyperplane intersects $t$-axis at $t = g(\lambda)$
**epigraph variation:** same interpretation if $G$ is replaced with

$$\mathcal{A} = \{(u, t) \mid f_1(x) \leq u, f_0(x) \leq t \text{ for some } x \in D\}$$

**strong duality**

- holds if there is a non-vertical supporting hyperplane to $\mathcal{A}$ at $(0, p^*)$
- for convex problem, $\mathcal{A}$ is convex, hence has supp. hyperplane at $(0, p^*)$
- Slater’s condition: if there exist $\tilde{u}, \tilde{t} \in \mathcal{A}$ with $\tilde{u} < 0$, then supporting hyperplanes at $(0, p^*)$ must be non-vertical
Complementary slackness

Assume strong duality holds, \( x^* \) is primal optimal, \((\lambda^*, \nu^*)\) is dual optimal

\[
f_0(x^*) = g(\lambda^*, \nu^*) = \inf_x \left( f_0(x) + \sum_{i=1}^{m} \lambda^*_i f_i(x) + \sum_{i=1}^{p} \nu^*_i h_i(x) \right)
\]

\[
\leq f_0(x^*) + \sum_{i=1}^{m} \lambda^*_i f_i(x^*) + \sum_{i=1}^{p} \nu^*_i h_i(x^*)
\]

\[
\leq f_0(x^*)
\]

Hence, the two inequalities hold with equality

- \( x^* \) minimizes \( L(x, \lambda^*, \nu^*) \)
- \( \lambda^*_i f_i(x^*) = 0 \) for \( i = 1, \ldots, m \) (known as complementary slackness):

\[
\lambda_i^* > 0 \implies f_i(x^*) = 0, \quad f_i(x^*) < 0 \implies \lambda_i^* = 0
\]
Karush-Kuhn-Tucker (KKT) conditions

the following four conditions are called KKT conditions (for a problem with differentiable $f_i, h_i$):

1. primal constraints: $f_i(x) \leq 0$, $i = 1, \ldots, m$, $h_i(x) = 0$, $i = 1, \ldots, p$
2. dual constraints: $\lambda \succeq 0$
3. complementary slackness: $\lambda_i f_i(x) = 0$, $i = 1, \ldots, m$
4. gradient of Lagrangian with respect to $x$ vanishes:

$$
\nabla f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + \sum_{i=1}^{p} \nu_i \nabla h_i(x) = 0
$$

from page 1–17: if strong duality holds and $x$, $\lambda$, $\nu$ are optimal, then they must satisfy the KKT conditions
KKT conditions for convex problem

if $\tilde{x}$, $\tilde{\lambda}$, $\tilde{\nu}$ satisfy KKT for a convex problem, then they are optimal:

- from complementary slackness: $f_0(\tilde{x}) = L(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$
- from 4th condition (and convexity): $g(\tilde{\lambda}, \tilde{\nu}) = L(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$

hence, $f_0(\tilde{x}) = g(\tilde{\lambda}, \tilde{\nu})$

if **Slater’s condition** is satisfied:

$x$ is optimal if and only if there exist $\lambda$, $\nu$ that satisfy KKT conditions

- recall that Slater implies strong duality, and dual optimum is attained
- generalizes optimality condition $\nabla f_0(x) = 0$ for unconstrained problem
example: water-filling (assume $\alpha_i > 0$)

$$\begin{align*}
\text{minimize} & \quad -\sum_{i=1}^{n} \log(x_i + \alpha_i) \\
\text{subject to} & \quad x \succeq 0, \quad 1^T x = 1
\end{align*}$$

$x$ is optimal iff $x \succeq 0$, $1^T x = 1$, and there exist $\lambda \in \mathbb{R}^n$, $\nu \in \mathbb{R}$ such that

$$\lambda \succeq 0, \quad \lambda_i x_i = 0, \quad \frac{1}{x_i + \alpha_i} + \lambda_i = \nu$$

- if $\nu < 1/\alpha_i$: $\lambda_i = 0$ and $x_i = 1/\nu - \alpha_i$
- if $\nu \geq 1/\alpha_i$: $\lambda_i = \nu - 1/\alpha_i$ and $x_i = 0$
- determine $\nu$ from $1^T x = \sum_{i=1}^{n} \max\{0, 1/\nu - \alpha_i\} = 1$

interpretation

- $n$ patches; level of patch $i$ is at height $\alpha_i$
- flood area with unit amount of water
- resulting level is $1/\nu^*$
Perturbation and sensitivity analysis

(unperturbed) optimization problem and its dual

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

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

perturbed problem and its dual

\[
\begin{align*}
\text{min.} & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq u_i, \quad i = 1, \ldots, m \\
& \quad h_i(x) = v_i, \quad i = 1, \ldots, p
\end{align*}
\]

\[
\begin{align*}
\text{max.} & \quad g(\lambda, \nu) - u^T \lambda - v^T \nu \\
\text{s.t.} & \quad \lambda \succeq 0
\end{align*}
\]

- $x$ is primal variable; $u$, $v$ are parameters
- $p^*(u, v)$ is optimal value as a function of $u$, $v$
- we are interested in information about $p^*(u, v)$ that we can obtain from the solution of the unperturbed problem and its dual
**global sensitivity result**

assume strong duality holds for unperturbed problem, and that $\lambda^*$, $\nu^*$ are dual optimal for unperturbed problem

apply weak duality to perturbed problem:

$$p^*(u, v) \geq g(\lambda^*, \nu^*) - u^T\lambda^* - v^T\nu^*$$

$$= p^*(0, 0) - u^T\lambda^* - v^T\nu^*$$

**sensitivity interpretation**

- if $\lambda_i^*$ large: $p^*$ increases greatly if we tighten constraint $i$ ($u_i < 0$)
- if $\lambda_i^*$ small: $p^*$ does not decrease much if we loosen constraint $i$ ($u_i > 0$)
- if $\nu_i^*$ large and positive: $p^*$ increases greatly if we take $v_i < 0$;
  if $\nu_i^*$ large and negative: $p^*$ increases greatly if we take $v_i > 0$
- if $\nu_i^*$ small and positive: $p^*$ does not decrease much if we take $v_i > 0$;
  if $\nu_i^*$ small and negative: $p^*$ does not decrease much if we take $v_i < 0$
**local sensitivity:** if (in addition) \( p^*(u, v) \) is differentiable at \((0, 0)\), then

\[
\lambda_i^* = -\frac{\partial p^*(0, 0)}{\partial u_i}, \quad \nu_i^* = -\frac{\partial p^*(0, 0)}{\partial v_i}
\]

proof (for \( \lambda_i^* \)): from global sensitivity result,

\[
\frac{\partial p^*(0, 0)}{\partial u_i} = \lim_{t \searrow 0} \frac{p^*(te_i, 0) - p^*(0, 0)}{t} \geq -\lambda_i^*
\]

\[
\frac{\partial p^*(0, 0)}{\partial u_i} = \lim_{t \nearrow 0} \frac{p^*(te_i, 0) - p^*(0, 0)}{t} \leq -\lambda_i^*
\]

hence, equality

\( p^*(u) \) for a problem with one (inequality) constraint:
Duality and problem reformulations

- equivalent formulations of a problem can lead to very different duals
- reformulating the primal problem can be useful when the dual is difficult to derive, or uninteresting

Common reformulations

- introduce new variables and equality constraints
- make explicit constraints implicit or vice-versa
- transform objective or constraint functions

\[ e.g., \text{replace } f_0(x) \text{ by } \phi(f_0(x)) \text{ with } \phi \text{ convex, increasing} \]
Introducing new variables and equality constraints

\[ \text{minimize } f_0(Ax + b) \]

- dual function is constant: \( g = \inf_x L(x) = \inf_x f_0(Ax + b) = p^* \)
- we have strong duality, but dual is quite useless

reformulated problem and its dual

\[ \begin{align*}
\text{minimize} & \quad f_0(y) \\
\text{subject to} & \quad Ax + b - y = 0
\end{align*} \quad \begin{align*}
\text{maximize} & \quad b^T \nu - f_0^*(\nu) \\
\text{subject to} & \quad A^T \nu = 0
\end{align*} \]

dual function follows from

\[
g(\nu) = \inf_{x,y} (f_0(y) - \nu^T y + \nu^T Ax + b^T \nu) \\
= \begin{cases} 
- f_0^*(\nu) + b^T \nu & A^T \nu = 0 \\
- \infty & \text{otherwise}
\end{cases}
\]
norm approximation problem: minimize $\|Ax - b\|$ 

minimize $\|y\|$ 
subject to $y = Ax - b$

can look up conjugate of $\| \cdot \|$, or derive dual directly

$$g(\nu) = \inf_{x,y}(\|y\| + \nu^T y - \nu^T Ax + b^T \nu)$$

$$= \begin{cases} 
  b^T \nu + \inf_y(\|y\| + \nu^T y) & A^T \nu = 0 \\
  -\infty & \text{otherwise}
\end{cases}$$

$$= \begin{cases} 
  b^T \nu & A^T \nu = 0, \quad \|\nu\|_* \leq 1 \\
  -\infty & \text{otherwise}
\end{cases}$$

(see page 1–4)

dual of norm approximation problem

maximize $b^T \nu$ 
subject to $A^T \nu = 0, \quad \|\nu\|_* \leq 1$
Implicit constraints

**LP with box constraints:** primal and dual problem

\[
\begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to} & \quad Ax = b \\
& \quad -1 \leq x \leq 1
\end{align*}
\]

\[
\begin{align*}
\text{maximize} & \quad -b^T \nu - 1^T \lambda_1 - 1^T \lambda_2 \\
\text{subject to} & \quad c + A^T \nu + \lambda_1 - \lambda_2 = 0 \\
& \quad \lambda_1 \geq 0, \quad \lambda_2 \geq 0
\end{align*}
\]

reformulation with box constraints made implicit

\[
\begin{align*}
\text{minimize} & \quad f_0(x) = \begin{cases} 
  c^T x & -1 \leq x \leq 1 \\
  \infty & \text{otherwise}
\end{cases} \\
\text{subject to} & \quad Ax = b
\end{align*}
\]

dual function

\[
g(\nu) = \inf_{-1 \leq x \leq 1} (c^T x + \nu^T (Ax - b)) = -b^T \nu - \| A^T \nu + c \|_1
\]

dual problem: maximize \(-b^T \nu - \| A^T \nu + c \|_1\)
Problems with generalized inequalities

minimize \( f_0(x) \)
subject to \( f_i(x) \leq_{K_i} 0, \quad i = 1, \ldots, m \)
\( h_i(x) = 0, \quad i = 1, \ldots, p \)

\( \leq_{K_i} \) is generalized inequality on \( \mathbb{R}^{k_i} \)

definitions are parallel to scalar case:

- Lagrange multiplier for \( f_i(x) \leq_{K_i} 0 \) is vector \( \lambda_i \in \mathbb{R}^{k_i} \)
- Lagrangian \( L : \mathbb{R}^n \times \mathbb{R}^{k_1} \times \cdots \times \mathbb{R}^{k_m} \times \mathbb{R}^p \rightarrow \mathbb{R} \), is defined as

\[
L(x, \lambda_1, \cdots, \lambda_m, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i^T f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)
\]

- dual function \( g : \mathbb{R}^{k_1} \times \cdots \times \mathbb{R}^{k_m} \times \mathbb{R}^p \rightarrow \mathbb{R} \), is defined as

\[
g(\lambda_1, \ldots, \lambda_m, \nu) = \inf_{x \in D} L(x, \lambda_1, \cdots, \lambda_m, \nu)
\]
**lower bound property:** if $\lambda_i \succeq K_i^* 0$, then $g(\lambda_1, \ldots, \lambda_m, \nu) \leq p^*$

**proof:** if $\tilde{x}$ is feasible and $\lambda \succeq K_i^* 0$, then

\[
f_0(\tilde{x}) \geq f_0(\tilde{x}) + \sum_{i=1}^{m} \lambda_i^T f_i(\tilde{x}) + \sum_{i=1}^{p} \nu_i h_i(\tilde{x})
\]

\[
\geq \inf_{x \in D} L(x, \lambda_1, \ldots, \lambda_m, \nu)
\]

\[
= g(\lambda_1, \ldots, \lambda_m, \nu)
\]

minimizing over all feasible $\tilde{x}$ gives $p^* \geq g(\lambda_1, \ldots, \lambda_m, \nu)$

**dual problem**

 maximize $\quad g(\lambda_1, \ldots, \lambda_m, \nu)$

 subject to $\quad \lambda_i \succeq K_i^* 0, \quad i = 1, \ldots, m$

• **weak duality:** $p^* \geq d^*$ always

• **strong duality:** $p^* = d^*$ for convex problem with constraint qualification
  (for example, Slater’s: primal problem is strictly feasible)
Semidefinite program

**primal SDP** \((F_i, G \in S^k)\)

\[
\begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to} & \quad x_1 F_1 + \cdots + x_n F_n \preceq G
\end{align*}
\]

- Lagrange multiplier is matrix \(Z \in S^k\)
- Lagrangian \(L(x, Z) = c^T x + \text{tr} \left( Z(x_1 F_1 + \cdots + x_n F_n - G) \right)\)
- dual function

\[
g(Z) = \inf_{x} L(x, Z) = \begin{cases} \quad -\text{tr}(GZ) & \text{tr}(F_i Z) + c_i = 0, \quad i = 1, \ldots, n \\
\quad -\infty & \text{otherwise} \end{cases}
\]

**dual SDP**

\[
\begin{align*}
\text{maximize} & \quad -\text{tr}(GZ) \\
\text{subject to} & \quad Z \succeq 0, \quad \text{tr}(F_i Z) + c_i = 0, \quad i = 1, \ldots, n
\end{align*}
\]

\(p^* = d^*\) if primal SDP is strictly feasible \((\exists x \text{ with } x_1 F_1 + \cdots + x_n F_n < G)\)