Proof of Strong duality under constraint qualification

$$\text{minimize} \quad f_0(x)$$
$$\text{s.t.} \quad f_i(x) \le 0, \quad i=1,\dots,m.$$
$$Ax = b. \qquad A = q \times n.$$

Suppose $f_0, \dots, f_m$ are convex.

Assume that there exists feasible $\tilde{x} \in \operatorname{relint} D$

Assume that D has nonempty interior. $\quad \operatorname{relint} D = \operatorname{int} D$

$\operatorname{rank}(A) = q$.

Suppose that p^* is finite.

$(0, 0, p^*) \in \mathcal{A}$.

$(u, 0, t \ge p^*) \in \mathcal{A} \quad u \ge 0$.

$u \to \infty$.

$t \to \infty$.

1°. $\mathcal{A} = \{ (u, v, t) \in \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{R} : \exists x \in D,$

$f_i(x) \le u_i, \ h_j(x) = v_j, \ f_0(x) \le t \}$ is convex.

2°. Define a convex set $\mathcal{B}$ as

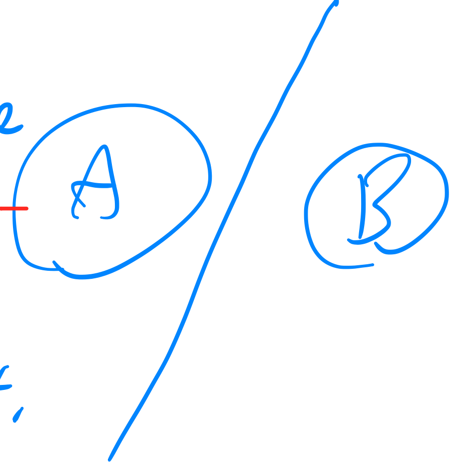
$$\mathcal{B} = \{ (0, 0, s) \in \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{R} : s < p^* \}$$

Claim: $\mathcal{A} \cap \mathcal{B} = \emptyset$. Let $(0, 0, t)$ with $t < p^*$ belong to $\mathcal{A}$.

$\exists x \in D$

$f_i(x) \le 0, \ h_j(x) = 0, \ f_0(x) \le t < p^*$.

3°. By the separating hyperplane theorem, there exists $(\tilde{\lambda}, \tilde{\nu}, \mu) \neq 0$, $\alpha \in \mathbb{R}$ s.t.



$$\begin{cases} (u,v,t) \in \mathcal{A} \Rightarrow \tilde{\lambda}^T u + \tilde{\nu}^T v + \mu t \geq \alpha & \text{--- (1)} \\ (u,v,t) \in \mathcal{B} \Rightarrow \tilde{\lambda}^T u + \tilde{\nu}^T v + \mu t \leq \alpha & \text{--- (2)} \end{cases}$$

(1) $\Rightarrow$ $\tilde{\lambda}^T u + \mu t$ is unbounded below over $\mathcal{A}$ unless $\tilde{\lambda} \geq 0$ and $\mu \geq 0$. $(u, 0, t)$

since $\left\langle \begin{matrix} u \\ t \end{matrix} \right.$ can go to $\left\{ \begin{matrix} +\infty \\ +\infty \end{matrix} \right.$.

(2) $(u,v,t) \in \mathcal{B} \Rightarrow u = 0,\ v = 0. \Rightarrow \mu t \leq \alpha$.

This is true for all $t < p^*$.

$\Rightarrow \mu(p^* - \epsilon) \leq \alpha.\ \forall \epsilon > 0$

$a - \epsilon \leq b\ \forall \epsilon > 0$

$\Rightarrow a \leq b$

$\Rightarrow \mu p^* \leq \alpha$.

Thus for $x \in D$, $(f_i(x), Ax-b, f_0(x)) \in \mathcal{A}$.

$\boxed{(1)}$ $\Rightarrow \sum_{i=1}^{m} \tilde{\lambda}_i f_i(x) + \tilde{\nu}^T (Ax-b) + \mu f_0(x) \geq \alpha$ —(3)

$\geq \mu p^*$.

4°. $\boxed{\text{Assume } \mu > 0}$

divide (3) by μ on both sides

$$\Rightarrow L(x, \tilde{\lambda}/\mu, \tilde{\nu}/\mu) \geq p^* \quad \forall x \in D.$$

Thus $g(\lambda, \nu) \geq p^*$, where $\lambda = \tilde{\lambda}/\mu$, $\nu = \tilde{\nu}/\mu$.

Also, the weak duality says that

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

$$\Rightarrow \underline{g(\lambda, \nu) = p^*.}$$

5°. $\boxed{\text{Assume } \mu = 0.}$

$$\forall x \in D. \quad \sum_{i=1}^m \tilde{\lambda}_i f_i(x) + \underline{\tilde{\nu}^T (Ax - b) \geq 0}$$

Consider $\tilde{x}$ with $A\tilde{x} - b = 0$, and $\underline{f_i(\tilde{x}) < 0}$

$$\Rightarrow \sum_{i=1}^m \tilde{\lambda}_i f_i(x) \geq 0. \Rightarrow \tilde{\lambda}_i = 0. \ \forall i.$$

$\Rightarrow$ Thus $\tilde{\lambda} = 0$.

From $(\underset{=0}{\tilde{\lambda}}, \tilde{\nu}, \underset{=0}{\mu}) \neq 0$. must have $\tilde{\nu} \neq 0$.

$$\tilde{\nu}^T (Ax - b) \geq 0$$

$\forall x \quad \tilde{\nu}^T (Ax - b) \geq 0.$

$\tilde{\nu}^T (A\tilde{x} - b) = 0.$

$\tilde{x} \in \operatorname{int} D$ by assumption

$\Downarrow$

$\exists B(\tilde{x}, r) \subseteq D.$ for $r > 0.$

$\exists x' \in B(\tilde{x}, r)$ s.t.

$\tilde{\nu}^T (Ax' - b) < 0.$

unless $\tilde{\nu}^T A = 0.$

$\therefore \quad \tilde{\nu}^T A = 0.$

$\tilde{\nu} \neq 0.$ A is full rank.

$\Rightarrow$ a contradiction.

$\Rightarrow \mu \neq 0.$

$$Ax = b.$$

If A is ~~not~~ full rank,

Ex:

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \Rightarrow \begin{array}{l} \boxed{x_1 + x_2 = b_1} \\ 0 = \underline{b_1 - b_2} \end{array}$$

$$E_1 \cdots E_k A x = E_1 \cdots E_k b,$$

where E_i are elementary matrices,

$$\Leftrightarrow \begin{bmatrix} A' \\ 0 \end{bmatrix} x = \begin{bmatrix} b' \\ ? \end{bmatrix}.$$

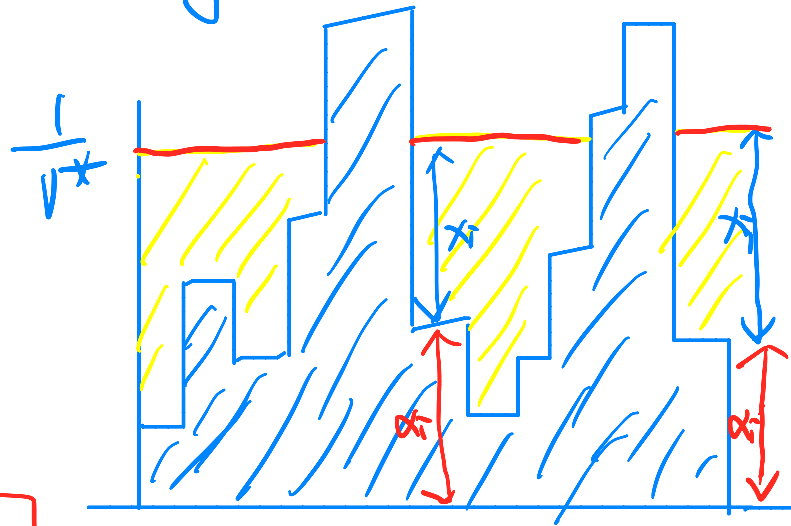
A' is full rank.

Example: Water-filling

minimize $-\sum_{i=1}^{n} \log(\alpha_i + x_i)$

s.t. $\underline{x \geq 0}$, $\sum_i x_i = 1$.

$-x \leq 0$.



where $\underline{\alpha_i > 0}$.
constant

The KKT conditions are

$x^* \geq 0$, $\sum_i x_i^* = 1$, $\lambda^* \geq 0$. $\boxed{\lambda_i^* x_i^* = 0}$ i

$$L(x, \lambda, \nu) = (\qquad) - \lambda^T x + \nu\left(\sum_i x_i - 1\right).$$

$$0 = \nabla L(\ ) \Rightarrow \frac{-1}{\alpha_i + x_i^*} - \lambda_i^* + \nu^* = 0.$$

$$\Rightarrow \nu^* - \frac{1}{\alpha_i + x_i^*} = \lambda_i^* \geq 0.$$

① $\lambda_i^* = 0$: $\nu^* = \frac{1}{\alpha_i + x_i^*} \Rightarrow x_i^* = \frac{1}{\nu^*} - \alpha_i$

$\leq \frac{1}{\alpha_i} \Rightarrow \left(\frac{1}{\nu^*} \geq \alpha_i\right)$

② $x_i^* = 0$: $\lambda_i^* = \nu^* - \frac{1}{\alpha_i} \Rightarrow \left(\frac{1}{\nu^*} \leq \alpha_i\right)$.

≥ 0.

determine $\boxed{\nu^*}$ from $1 = \sum_i x_i^* = \sum_{i=1}^{n} \max\left\{0, \frac{1}{\nu^*} - \alpha_i\right\}$.

Duality and problem reformulation.

$$\text{minimize} \quad \|Ax-b\|$$

$$L(x) = \|Ax-b\|$$

$$g(\ ) = \inf_x \|Ax-b\| = p^*.$$

Reformulating the problem can be useful when the dual is difficult or uninteresting.

1. Introduce new variables and associated equality constraints.
2. Making explicit constraint implicit or vice-versa.
3. Transforming the objective function or constraint functions.

eg. $f_0(x)$ convex.

$\rightarrow \phi(f_0(x))$ where ϕ is convex & increasing.

(composition rule).

minimize $\|y\| = f_0(y)$. $\iff$ minimize $\|Ax-b\|$

s.t. $Ax-b=y$.

variables x, y.

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

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

(otherwise unbounded)

$$= \begin{cases} -b^T\nu - \underline{\underline{f_0^*(\nu)}}, & \text{if } \nu^T A = 0. \\ -\infty, & \text{otherwise}. \end{cases}$$

$\boxed{\text{conjugate function}}$

$$f^*(x) = \sup_y \{ y^T x - f(x) \}.$$

$$-f^*(x) = \inf_y \{ f(x) - y^T x \}.$$

$\boxed{\text{dual norm.}}$

P. 93. $f_0(y) = \|y\|$

$$\Rightarrow f_0^*(\nu) = \begin{cases} 0, & \text{if } \|\nu\|_* \le 1 \\ \infty, & \text{otherwise}. \end{cases}$$

minimize $b^T\nu$

s.t. $A^T\nu = 0$, $\|\nu\|_* \le 1$.

Problems with generalized inequalities

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \preceq_{K_i} 0, \quad i=1,\dots,m \\ & h_j(x) = 0, \quad j=1,\dots,q. \end{aligned}$$

where $\preceq_{K_i}$ is the generalized inequality on $\mathbb{R}^{k_i}$

Ex. $K_i = \mathbb{R}^n_+$. S^n_+.

— Definitions are parallel to those in a problem with scalar inequalities.

— Lagrange multiplier for $f_i(x) \preceq_{K_i} 0$ is a vector $\lambda_i \in \mathbb{R}^{k_i}$.

Lagrangian $L: \mathbb{R}^n \times \mathbb{R}^{k_1} \times \cdots \times \mathbb{R}^{k_m} \times \mathbb{R}^p$ defined by

$$L(x, \lambda_1, \dots, \lambda_m, \nu) = f_0(x) + \sum_{i=1}^{m} \underbrace{\lambda_i^T f_i(x)}_{\text{inner product of } \lambda_i \,\&\, f_i(x)} + \sum_{j=1}^{q} \nu_j h_j(x).$$

dual function $g: \mathbb{R}^{k_1} \times \cdots \times \mathbb{R}^{k_m} \times \mathbb{R}^{q} \to \mathbb{R}$.

$$g(\lambda_1, \ldots, \lambda_m, \nu) = \inf_{x \in D} L(x, \lambda_1, \ldots, \lambda_m, \nu).$$

· The nonnegativity requirement on the dual variable is replaced by the condition

$$\lambda_i \succeq_{K_i^*} 0, \quad i = 1, \ldots, m.$$

$$f_i(x) \preceq_{K_i} 0 \Rightarrow 0 - f_i(x) \in K_i$$

$$K^* = \{y : y^T x \geq 0 \ \forall x \in K\}$$

$$\lambda_i \succeq_{K_i^*} 0 \Rightarrow \lambda_i \in K_i^*$$

$$-\lambda_i^T f_i(x) \geq 0 \text{ or } \lambda_i^T f_i(x) \leq 0.$$

(usually consider $\mathbb{R}_+^n$, S_+^n self-dual cones)

(In the scalar case, $\lambda_i \geq 0$, $f_i(x) \leq 0$ $\Rightarrow \lambda_i f_i(x) \leq 0$)

Weak duality. , $\tilde{x}$ feasible.

$$f_0(\tilde{x}) + \sum_{i=1}^{m} \underbrace{\lambda_i^T f_i(\tilde{x})}_{\leq 0.} + \sum_{j=1}^{q} \underbrace{\nu_j h_j(\tilde{x})}_{= 0}$$

$$\leq f_0(\tilde{x}).$$

$$g(\lambda_1, \ldots, \lambda_m, \nu) \leq p^*.$$

- dual problem

$$\begin{aligned} &\text{maximize} && g(\lambda_1, \ldots, \lambda_m, \nu) \\ &\text{s.t.} && \lambda_i \succeq_{K_i^*} 0, \quad i = 1, \ldots, m. \end{aligned}$$

- Weak duality: $d^* \leq p^*$.

- Slater's condition: Strong duality holds ($d^* = p^*$) if there exists an $x \in \operatorname{relint} D$ with $\underline{Ax = b}$ and $\boxed{f_i(x) \prec_{K_i} 0}$ $i = 1, \ldots, m$

affine equality constraint.

Semidefinite program (SDP)

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

$$\text{minimize} \quad c^T x$$
$$\text{s.t.} \quad \underbrace{x_1 F_1 + x_2 F_2 + \cdots + x_n F_n \preceq G}_{\text{linear matrix inequality}}.$$

variable x.

- one inequality constraint
  $K_1 = S_+^k$ the positive semidefinite cone.
- The Lagrange multiplier is a matrix $Z \in S^k$.
- Lagrangian

$$L(x, Z) = c^T x + \langle Z, x_1 F_1 + \cdots + x_n F_n - G \rangle$$
$$= c^T x + \mathrm{Tr}\left(Z\left(x_1 F_1 + \cdots + x_n F_n - G\right)\right)$$
$$= x_1 (c_1 + \mathrm{tr}\, Z F_1) + x_2 (c_2 + \mathrm{tr}\, Z F_2) + \cdots + x_n (c_n + \mathrm{tr}\, Z F_n) - \mathrm{tr}\, G Z$$

dual funtion

$$g(Z) = \inf_x L(x, Z).$$

$$= \begin{cases} -\operatorname{tr} GZ, & \text{if } \operatorname{tr}(F_i Z) + c_i = 0 \\ & i = 1, \ldots, m \\ -\infty, & \text{otherwise.} \end{cases}$$

dual SDP.

$$\begin{aligned} &\text{maximize} && -\operatorname{tr} GZ \\ &\text{s.t.} && Z \succeq 0 \\ & && \operatorname{tr} F_i Z + c_i = 0, \ i = 1, \ldots, n. \end{aligned}$$

- $p^* = d^*$ if the primal SDP is strictly feasible (i.e. $\exists x$ with $x_1 F_1 + \cdots + x_n F_n \prec G$).