(Review)

For convex opt. problem (p* = value of primal)

$$p^* = \min_x f(x) \quad \text{s.t.} \quad g_i(x) \le 0 \quad 1 \le i \le k, \qquad (1)$$

KKT conditions are

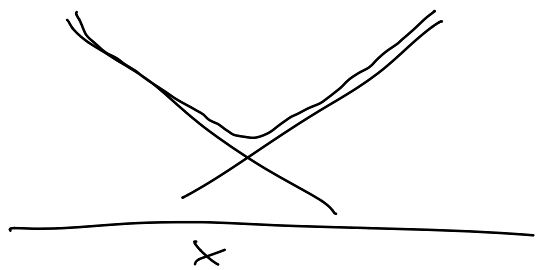
1. $g_i(x^*) \le 0$ for $1 \le i \le k$
2. $\lambda^*(i) \ge 0$ for $1 \le i \le k$
3. $\lambda^*(i) g_i(x^*) = 0$ for $1 \le i \le k$
4. $\nabla f(x^*) + \sum_i \lambda^*(i) \nabla g_i(x^*) = 0$

Thm 1 If $f, g_1 \dots g_k$ are convex and differentiable and x^*, λ^* satisfy 1-4, then x^* is optimal. If the g_i are linear or if exists strictly feasible x_0 ($g_i(x_0) < 0 \ \forall i$) then there exist x^* and λ^* that satisfy 1-4.

Can we replace $g_1, \dots g_k$ with

$$G(x) = \max_i g_i(x) \quad ?$$

- $G(x) \le 0 \iff g_i(x) \le 0, \ \forall i$
- G is convex because max of convex is convex



- max is not differentiable $\times$

$\max(x_1, x_2)$ at $(1,1)$

$t = (1,0)$

$$\max(x + \varepsilon t) = \begin{cases} 1+\varepsilon & \varepsilon > 0 \\ 1 & \varepsilon < 0 \end{cases}$$

If had $\min f(x)$ s.t. $G(x) \le 0$

$$\Downarrow$$

$\min f(x)$ s.t. $g_i(x) \le 0 \quad 1 \le i \le k$

Lagrange $\quad L(x, \lambda) = f(x) + \sum_i \lambda(i) g_i(x)$

$$q(\lambda) = \inf_x L(x, \lambda)$$

$$d^* = \max_{\lambda \ge 0} q(\lambda)$$

Lem 1 $\quad q(\lambda) \le f(x) \quad \forall \lambda \ge 0$, feasible x (weak duality)

So $d^* \le p^*$

proof $\quad x$ feasible $\Leftrightarrow g_i(x) \le 0, \forall i$

$$\lambda \ge 0 \Rightarrow \sum_i \lambda(i) g_i(x) \le 0$$

$$\Rightarrow L(x, \lambda) = f(x) + \sum_i \lambda(i) g_i(x) \le f(x)$$

$$q(\lambda) = \inf_{x_0} L(x_0, \lambda) \le L(x, \lambda)$$

$$q(\lambda) \le f(x)$$

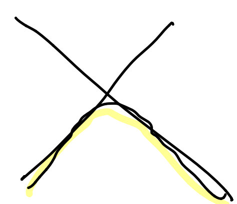
$q(\lambda) = \inf_x L(x, \lambda) \qquad$ inf can give $-\infty$

no constraints on x

$\max_\lambda q(\lambda) \qquad q(\lambda)$ is concave

$L(x, \lambda)$ for every x is linear in λ

inf of linear fncs is concave



Thm 1 (strong duality)

If $f, g_1 \ldots g_d$ are differentiable & convex

and either $g_1 \ldots g_d$ are linear or strictly feasible

then $d^* = p^*$

<u>Proof</u> KKT $\to x^*, \lambda^*$ $\quad x^*$ feas, $\lambda^* \geq 0$

let $h(x) = L(x, \lambda^*) = f(x) + \sum_i \underline{\lambda^*(i)} g_i(x)$

g_i convex, $\lambda^*(i) \geq 0$

$\Rightarrow h(x)$ is convex

$\nabla f(x^*) + \sum_i \lambda^*(i) \nabla g_i(x^*) = 0 = \nabla h(x^*)$

$\Rightarrow x^*$ is a global min of h

$q(\lambda^*) = \inf_x L(x, \lambda^*) = \inf_x h(x) = h(x^*)$

KKT3 $\Rightarrow \sum_i \lambda^*(i) g_i(x) = 0$

$h(x^*) = f(x^*) = q(\lambda^*)$

$d^* \geq p^* \Rightarrow d^* = p^*$

For general feasible x, $\lambda \geq 0$

$f(x) \geq q(\lambda)$ $\qquad f(x) - q(\lambda) = p - d$

"duality gap"

$h(x) = 0 \qquad h(x) \in \mathbb{R} \qquad h(x) \leq 0 \qquad -h(x) \leq 0$

$\qquad\qquad\qquad\qquad\qquad\qquad \lambda_+ \qquad\qquad \lambda_-$

$\lambda_+ h(x) + \lambda_-(-h(x)) = (\lambda_+ - \lambda_-) h(x)$

$\lambda_+, \lambda_- \geq 0$

$\nu = \lambda_+ - \lambda_- \qquad = \qquad \nu h(x)$

$$\min_x \; f(x) \quad \text{s.t.} \quad g_i(x) \le 0 \quad 1 \le i \le d$$
$$h_i(x) = 0 \quad 1 \le i \le c$$

$$L(x, \lambda, v) = f(x) + \sum_{i=1}^{d} \lambda(i)\, g_i(x) + \sum_{i=1}^{c} v(i)\, h_i(x)$$

$$q(\lambda, v) = \inf_x L(x, \lambda, v)$$

$$d^* = \max \; q(\lambda, v) \quad \lambda \in \mathbb{R}^d_+ \quad v \in \mathbb{R}^c$$

LP $\quad \max_x \; c^T x \quad \text{s.t.} \quad a_i^T x \le b_i \quad 1 \le i \le d$

$x \in \mathbb{R}^n$

replace c by $(-c)$

$$\min_x \; c^T x \quad \text{s.t.} \quad a_i^T x \le b_i$$
$$g_i(x) = 0 \qquad g_i(x) = a_i^T x - b_i \le 0$$

$$L(x, \lambda) = c^T x + \sum_i \lambda_i g_i(x) = c^T x + \sum_i \lambda_i (a_i^T x - b_i)$$

$$= c^T x + \lambda^T A x - \lambda^T b$$
$$= (c^T + \lambda^T A) x - \lambda^T b$$

$$q(\lambda) = \inf_x (c^T + \lambda^T A) x - \lambda^T b$$

if $c^T + \lambda^T A \neq 0$, $q(\lambda) = -\infty$

$$q(\lambda) = \begin{cases} -\lambda^T b & \text{if } c^T + \lambda^T A = \bar{0} \\ -\infty & \text{o.w.} \end{cases}$$

$$\max_{\lambda \geq 0} q(\lambda) = \max_{\lambda \geq 0} -\lambda^T b \quad \text{s.t. } \lambda^T A = -c^T$$

$$\Updownarrow$$

$$A^T \lambda = -c$$

$$\min \; c^T x \quad \text{s.t. } Ax = b, \; x \geq 0$$

$$g_i(x) = -x_i \leq 0$$

$$L(x, \lambda, \nu) = c^T x + \sum_i \lambda(i)(-x(i)) + \sum_i \nu(i)(a_i^T x - b_i)$$

$$= c^T x - \lambda^T x + \nu^T (Ax - b)$$

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

$$q(\lambda, \nu) = \inf_x (c - \lambda + A^T \nu)^T x - \nu^T b$$

$$= -\infty \text{ unless } c - \lambda + A^T \nu = 0$$

eliminate λ $\quad \exists \lambda \geq 0$ s.t. $c - \lambda + A^T \nu = 0$

iff $c + A^T \nu \geq 0$

$$\max_{\substack{\lambda, \nu \\ \lambda \geq 0}} q(\lambda, \nu) = \max_{\nu} -b^T\nu \quad \text{s.t.} \quad A^T\nu + c \geq 0$$

$$\min \|x\|_2^2 \quad \text{s.t.} \quad Ax = b$$

x^Tx

$$L(x, \nu) = \|x\|_2^2 + \nu^T(Ax - b) = \|x\|_2^2 + \nu^TAx - \nu^Tb$$

$$q(\nu) = \inf_x L(x, \nu)$$

$$\nabla_x L(x, \nu) = 2x + A^T\nu = 0$$

$$x = -\frac{1}{2}A^T\nu$$

$$q(\nu) = \frac{1}{4}\nu^TAA^T\nu - \frac{1}{2}\nu^TAA^T\nu - \nu^Tb$$

$$= -\frac{1}{4}\nu^TAA^T\nu - \nu^Tb$$

$\min \|x\|$ s.t. $Ax = b$, $\quad \|\cdot\|$ arbitrary

Dual norm $\|\cdot\|_*$

$$\|y\|_* = \max_x x^Ty \quad \|x\| \leq 1$$

will be achieved $\|x\| = 1$

dual of $\|\cdot\|_2$ is $\|\cdot\|_2$

Cauchy-Schwartz $\quad x^T y \le \|x\|_2 \, \|y\|_2 \le \|y\|_2$

$x = y / \|y\|_2 \qquad \frac{y^T}{\|y\|_2} \cdot y = \|y\|_2$

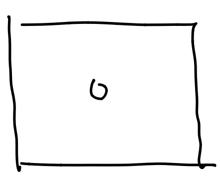
dual $\|\cdot\|_\infty$ is $\|\cdot\|_1 \qquad \|x\|_\infty = \max_i |x(i)|$

given y set. $x(i) = \begin{cases} 1 & \text{if } y(i) \ge 0 \\ -1 & \text{if } y(i) < 0 \end{cases}$

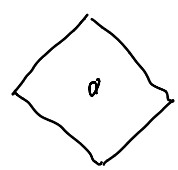
$\|x\|_\infty = 1$

$x^T y = \sum_i x(i)\, y(i) = \sum_i |y(i)| = \|y\|_1$

$\|\cdot\|_{**} = \|\cdot\| \implies$ dual of $\|\cdot\|_1$ is $\|\cdot\|_\infty$



$\|x\|_\infty \le 1$



$\|x\|_1 \le 1$

$\|x\|_p = \left(\sum_i |x(i)|^p \right)^{1/p} \qquad \|x\|_p$ and $\|x\|_q$ are dual when $\frac{1}{p} + \frac{1}{q} = 1$

$1 \le p, q \le \infty$

Hölder $\quad x^T y \le \|x\|_p \, \|y\|_q$ when $\frac{1}{p} + \frac{1}{q} = 1$

tight $\quad y(i)^q = c \cdot x(i)^p \qquad y, x \ge 0 \quad$ any $c > 0$

$$\boxed{\min \|x\| \text{ s.t. } Ax = b}$$

$$g(\nu) = \inf_x \|x\| + \nu^T A x - \nu^T b$$

if $\|\nu^T A\|_* \le 1$ then $\forall x$ $\|x\| \ge |\nu^T A x|$

$$\Rightarrow \|x\| + \nu^T A x \ge 0$$

inf when $x = 0$. get $-\nu^T b$

if $\|\nu^T A\|_* > 1$ $\exists u$ s.t. $\|u\| = 1$

$$\nu^T A u = \|\nu^T A\|_* > 1$$

$x = cu$ $\quad c \to -\infty$

$\|x\| = c$ $\quad \nu^T A x = c\|\nu^T A\|_* < -c$

$\inf \to -\infty$

$$g(\nu) = \begin{cases} -b^T\nu & \text{if } \|\nu^T A\|_* \le 1 \\ -\infty & \text{o.w.} \end{cases}$$

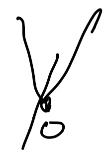
$\max_\nu g(\nu)$ $\quad \boxed{\max_\nu -b^T\nu \text{ s.t. } \|A^T\nu\|_* \le 1}$

$g_i(x) \le 0$ $\quad g$ diff & convex

PSD matrices $\quad M : x^T M x \ge 0 \quad \forall x$

S^n = n-by-n sym. $\quad S^n_+$ pos def

K is cone if $x \in K \Rightarrow tx \in K \quad \forall t \geq 0$



proper if

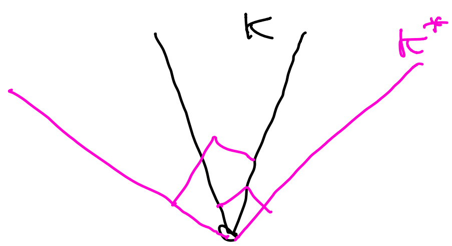
a. convex

b. closed

c. solid: have an interior

d. pointed $x \in K,\ x \neq 0,\ \Rightarrow -x \notin K$

The dual cone is $K^* = \{x : x^T y \geq 0,\ \forall y \in K\}$



$K^{**} = K$ in finite dim

$\mathbb{R}^n_+$ dual $\mathbb{R}^n_+$

S^n_+ dual is S^n_+

$X, Y \in S^n$ $\quad \cancel{X^T Y}$ treat X and Y as vectors

$\langle X, Y \rangle = \mathrm{Trace}(X^T Y)$

$X \preceq_K Y$ iff $Y - X \in K$

$0 \preceq_K X$ iff $X \in K$

$-X \preceq_K 0$ iff $X \in K$

$$\min_x f(x) \quad \text{s.t.} \quad g_i(x) \le 0 \qquad 1 \le i \le d$$
$$-X \preceq_{K_i} 0 \qquad 1 \le i \le c$$

Lagrage dual:

$$L(x, \lambda_0, \lambda_1 .. \lambda_c) = f(x) + \sum_{i=1}^{d} \lambda_0(i)\, g_i(x) + \sum_{i=1}^{c} \langle \lambda_i, -x \rangle$$