Lec 9: KKT, Perturbation

Alternatives

- Complementary slackness

Assume strong duality holds, x^* is primal optima. (λ^*, ν^*) is dual optimal

$$f_0(x^*) = g(\lambda^*, \nu^*) = \inf_{x \in D} \left[f_0(x) + \sum_i \lambda_i f_i(x) + \sum_i \nu_i h_i(x) \right]$$

$$\leq f_0(x^*) + \sum_i \lambda_i f_i(x^*) + \sum_i \nu_i h_i(x^*)$$

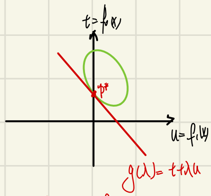
$$\leq f_0(x^*)$$

hence the inequalites hold with equality

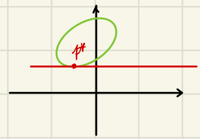
x^* minimizes the lagrangian

$\lambda_i^* f_i(x^*) = 0$ (complementary slackness)

$\lambda_i^* > 0 \Rightarrow f_i(x^*) = 0$ $\qquad$ $f_i(x^*) < 0 \quad \lambda_i^* = 0$



$\lambda > 0, \ f_1(x) = 0$



$f_1(x) < 0 \quad \lambda = 0$

- Karush - Kuhn - Tucker (KKT) conditions
  (for a problem with differentiable f_i h_i)

1. KKT condition

   primal feasibility: $f_i(x) \leq 0 \quad h_i(x) = 0$

   dual feasibility: $\lambda_i \geq 0$

   complementary slackness: $f_i(x) \cdot \lambda_i = 0$

   Gradient of L w.r.t x vanishes:
   $$\nabla_x L = \nabla f_0(x) + \sum_i \lambda_i f_i(x) + \sum_i \nu_i h_i(x) = 0$$

   if strong duality holds for a problem (convex or not)
   the KKT conditions hold for a primal-optimal-dual-optimal pair.

2. KKT condition for convex problems

if $\tilde{x}, \tilde{\lambda}, \tilde{\nu}$ satisfies the KKT condition $\Leftrightarrow$ then they are optimal

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

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

$\begin{cases} \text{for any problem, convex or not, strong duality} \Rightarrow \text{optimal satisfies KKT condition} \\ \text{for a convex problem, if slater condition holds, KKT condition} \Leftrightarrow \text{optimal} \end{cases}$

- Perturbation and sensitivity analysis

(unperturbed) optimization and its dual

$$\begin{array}{lll} \min. & f_0(x) & \\ \text{s.t.} & f_i(x) \le 0 & \\ & h_i(x) = 0 & \end{array} \qquad \begin{array}{ll} \max. & g(\lambda, \nu) \\ \text{s.t.} & \lambda \succeq 0 \end{array}$$

perturbed optimization and its dual (relaxed constraints)

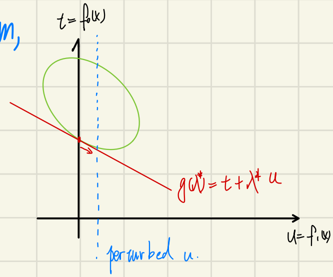
$$\begin{array}{ll} \min & f_0(x) \\ \text{s.t.} & f_i(x) \le u_i \\ & h_i(x) = v_i \end{array} \qquad \max. \ g(\lambda, \nu) - \lambda^T u - \nu^T v$$

(loosening and tightening)
(shifting i.e. hit a target)

x is the primal variable, u and v are parameter.
$p^*(u, v)$ is optimal value as a function of u, v
what can we say about $p^*(u, v)$ from the unperturbed problem and its dual

1. global sensitivity result:

assume strong duality holds for the unperturbed problem, and λ^*, ν^* is dual optimal of unperturbed problem



apply weak duality to perturbated problem:

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

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

2. Sensitivity interpretation:

if λ_i^* large: p^* increase greatly if we tighten constraint i $(u_i < 0)$
(p^* becomes inf if perturbated problem is infeasible)

if λ_i^* small: p^* does not decrease much if we loosen constraint i $(u_i > 0)$

3. local sensitivity:

if $p^*(u,v)$ is differentiable at $(0,0)$, then

$$\lambda_i^* = -\frac{\partial p^*(0,0)}{\partial u_i} \qquad \nu_i = -\frac{\partial p^*(0,0)}{\partial v_i}$$

eg. $\min f(x)$
s.t. $f_1(x) \leq 0$
$\Downarrow$
$f_1(x^*) = -0.1 \quad \lambda_1^* = 0$
we can perturb the problem $f_1(x) \leq 0.1$
p^* won't change

$$\begin{cases} \lim\limits_{t \to 0} \dfrac{p^*(t \cdot e_i, 0)}{t} = \dfrac{\partial p^*(0,0)}{\partial u_i} \\ p^*(t \cdot e_i, 0) - p^*(0,0) \geq -\lambda_i^* \cdot t \quad \text{(global sensitivity)} \end{cases}$$

for $t > 0$ $\quad \dfrac{p^*(t \cdot e_i, 0) - p^*(0,0)}{t} \geq -\lambda_i^*$

$$\lim_{t \to 0} \frac{p^*(t \cdot e_i, 0) - p^*(0,0)}{t} = \frac{\partial p^*(0,0)}{\partial u_i} \geq -\lambda_i$$

for $t < 0$ $\quad \dfrac{p^*(t \cdot e_i, 0) - p^*(0,0)}{t} \leq -\lambda_i^*$

$$\lim_{t \to 0} \frac{p^*(t \cdot e_i, 0) - p^*(0,0)}{t} = \frac{\partial p^*(0,0)}{\partial u_i} \leq -\lambda_i$$

$$\frac{\partial p^*(0,0)}{\partial u_i} = \lambda_i^*$$

- Introducing new variables and equality constraints

$$\min.\ f_0(Ax+b)$$

dual function $g(\lambda) = \min_x f_0(Ax+b) = p^*$ (strong duality, but dual is useless)

reformulated problem and its dual

$$\begin{aligned} \min.\ & f_0(y) \\ \text{s.t.}\ & Ax+b-y=0 \end{aligned}$$

$$g(v) = \inf_{x,y}\left[f_0(y) - v^T y + v^T A x + b^T v \right]$$

$$= -\sup_{x,y}\left[-f_0(y) + v^T y - v^T A x - b^T v \right]$$

$$= \begin{cases} -f_0^*(v) + b^T v & A^T v = 0 \\ -\infty & A^T v \neq 0 \end{cases} \Longrightarrow \begin{aligned} \max.\ & b^T v - f_0^*(v) \\ \text{s.t.}\ & A^T v = 0 \end{aligned}$$

eg. $\min.\ \|Ax-b\|$ ($\|\cdot\|$ is any norm)

$$\begin{aligned} \min.\ & \|y\| \\ \text{s.t.}\ & y = Ax-b \end{aligned}$$

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

$$= \inf_{x,y}\left(\|y\| + v^T y - v^T A x \right) + b^T v$$

$$= \begin{cases} -\infty & A^T v \neq 0 \quad (\text{take } y=\vec{0},\ v^T(Ax) > 0 \text{ since } A^T v \neq 0 \text{ let } \|x\| \to \infty) \\ \inf_y \left(\|y\| + v^T y \right) + b^T v = \begin{cases} b^T v & \|v\|_* \leq 1 \\ -\infty & \text{otherwise} \end{cases} \end{cases}$$

$$\Downarrow$$

$$\begin{aligned} \max.\ & b^T v \\ \text{s.t.}\ & A^T v = 0 \\ & \|v\|_* \leq 1 \end{aligned}$$

- Implicit constraints.

eg. LP with box constraints

$$\begin{aligned} \min.\ & c^Tx \\ \text{s.t.}\ & Ax=b \\ & -1\le x\le 1 \end{aligned}$$

$$L(x,\lambda_1,\lambda_2,\nu) = c^Tx+\nu^T(Ax-b)+\lambda_1^T(x-1)+\lambda_2^T(x-1)$$
$$= (c+A^T\nu+\lambda_1-\lambda_2)^Tx - b^T\nu + 1^T\lambda_1 - 1^T\lambda_2$$

$$g(\lambda_1,\lambda_2,\nu) = \inf_x L(x,\lambda_1,\lambda_2,\nu) = \begin{cases} -b^T\nu+1^T\lambda_1-1^T\lambda_2 & \text{if } c+A^T\nu+\lambda_1-\lambda_2=0 \\ -\infty & \text{otherwise} \end{cases}$$

$\Downarrow$

$$\begin{aligned} \max.\ & -b^T\nu + 1^T\lambda_1 - 1^T\lambda_2 \\ \text{s.t.}\ & c+A^T\nu+\lambda_1-\lambda_2=0 \\ & \lambda_1\ge 0 \quad \lambda_2\ge 0 \end{aligned}$$

$$\min.\ f_0(x) = \begin{cases} c^Tx & -1\le x\le 1 \\ \infty & \text{otherwise} \end{cases}$$
$$\text{s.t. } Ax=b$$

$$g(x,\nu) = \inf_{-1\le x\le 1}\left[c^Tx + \nu^T(Ax-b)\right]$$
$$= \inf_{-1\le x\le 1}\left[(c+A^T\nu)^Tx\right] - b^T\nu$$
$$= -\|c+A^T\nu\|_1 - b^T\nu$$

$\Downarrow$

$$\max.\ -\|c+A^T\nu\|_1 - b^T\nu$$
$$(\ \|x\|_\infty \le 1)$$

- Theorem of alternatives

1. weak alternatives for nonstrict equalities

determine the feasibility of $f_i(x)\le 0 \quad i=1,\cdots,m \quad h_i(x)=0 \quad i=1,\cdots,p$

$D = \bigcap \operatorname{dom} f_i \bigcap \operatorname{dom} h_i$

$$\begin{aligned} \min.\ & 0 \\ \text{s.t.}\ & f_i(x)\le 0 \\ & h_i(x)=0 \end{aligned}$$

the problem has optimal value $p^* = \begin{cases} 0 & \text{if feasible} \\ -\infty & \text{if infeasible} \end{cases}$

$$g(\lambda,\nu) = \inf_x\left(\sum_i \lambda_i f_i(x) + \sum_i \nu_i h_i(x)\right)$$

$g(\lambda,\nu)$ is homogeneous in λ,ν

$$g(\alpha\lambda,\alpha\nu) = \alpha g(\lambda,\nu) \text{ for } \alpha>0$$

dual problem:
$$\begin{aligned} \max.\ & \inf_x\left(\sum_i \lambda_i f_i(x) + \sum_i \nu_i h_i(x)\right) \\ \text{s.t.}\ & \lambda_i\ge 0 \end{aligned}$$

$$d^* = \begin{cases} \infty & \text{if } \lambda\ge 0\ g(\lambda,\nu)>0 \text{ is feasible} \\ 0 & \text{if } \lambda\ge 0\ g(\lambda,\nu)>0 \text{ is infeasible} \end{cases}$$

if $\lambda \geq 0$ $g(\lambda, \nu) > 0$ is feasible

$d^* = \infty$ $p^* = \infty$; the original problem is infeasible

if the original problem is feasible

$\exists \tilde{x} \in D$ s.t. $f_i(\tilde{x}) \leq 0$ $h_i(\tilde{x}) = 0$

$g(\lambda, \nu) = \inf_x (L(x, \lambda, \nu)) \leq L(\tilde{x}, \lambda, \nu) \leq 0$

$\therefore$ $g(\lambda, \nu) \leq 0$

$$\left\{ \begin{array}{ll} f_i(x) \leq 0 & h_i(x) = 0 \\ \lambda \geq 0 & g(\lambda, \nu) > 0 \end{array} \right] \text{ at most one of them is feasible}$$

2. weak alternatives for strict inequalities

min 0

s.t. $f_i(x) < 0$

$h_i(x) = 0$

if the original problem is feasible

$\exists \tilde{x} \in D$ s.t. $f_i(\tilde{x}) < 0$ $h_i(\tilde{x}) = 0$

$g(\lambda, \nu) = \inf_x \left(\sum_i \lambda_i f_i(x) + \sum_i \nu_i h_i(x) \right)$

$\leq \sum_i \lambda_i f_i(\tilde{x}) + \sum_i \nu_i h_i(\tilde{x})$

< 0

$\therefore$ $g(\lambda, \nu) \geq 0$ $\lambda_i \geq 0$ $\lambda \neq 0$ is infeasible

$$\left\{ \begin{array}{l} f_i(x) < 0 \quad h_i(x) = 0 \\ g(\lambda, \nu) \geq 0 \quad \lambda_i \geq 0 \quad \lambda \neq 0 \end{array} \right\} \text{ at most one of them is feasible}$$

3. Strong alternatives for strict inequality

$f_i(x) < 0$ $f_i(x)$ are convex

$Ax = b$ $\exists x \in \operatorname{relint} D$ with $Ax = b$

$\Downarrow$

min. s

s.t. $f_i(x) - s \leq 0$

$Ax = b$

($p^* < 0$ iff the strict inequality system is feasible)

$L(x, \lambda, \nu) = s + \sum_i \lambda_i (f_i(x) - s) + \nu^T (Ax - b)$

$\inf_{x \in D, s} L(x, \lambda, \nu) = \inf_{x \in D, s} \left[(1 - \sum_i \lambda_i) s + \sum_i \lambda_i f_i(x) + \nu^T (Ax - b) \right]$

$$= \begin{cases} g(\lambda, \nu) & \mathbf{1}^T \lambda = 1 \\ -\infty & \text{otherwise} \end{cases}$$

$g(\lambda, \nu)$ is the dual function of $f_i \leq 0$, $Ax = b$

$\max.\ g(\lambda, \nu)$
s.t. $\mathbf{1}^T\lambda = 1$
$\lambda \succeq 0$

the slater's condition holds for

$\max.\ s$
s.t. $f_i(x) - s \le 0$
$Ax = b$

$\exists \tilde{x} \in \text{relint } D$ with $A\tilde{x} = b$
choosing any $\tilde{s} = \max_i f_i(\tilde{x})$
$(\tilde{x}, \tilde{s})$ is a strictly feasible point for the problem
$\therefore p^* = d^*$

$\Downarrow\ p^* = d^*$

$\exists (\lambda^*, \nu^*)$ such that
$g(\lambda^*, \nu^*) = p^*$ $\lambda^* \succeq 0$ $\mathbf{1}^T\lambda^* = 1$

1° suppose $f_i(x) < 0$ $Ax = b$ is infeasible
$p^* \ge 0$
(λ^*, ν^*) satisfies $\lambda \succeq 0$ $\lambda \ne 0$ $g(\lambda, \nu) \ge 0$

2° suppose $\lambda \succeq 0$ $\lambda \ne 0$ $g(\lambda, \nu) \ge 0$ is feasible
$p^* = d^* = \{\max.\ g(\lambda, \nu)$ s.t. $\lambda \succeq 0$ $\mathbf{1}^T\lambda = 1\} \ge 0$ ($g(\lambda, \nu)$ homogeneous in λ, ν)
$\therefore f_i(x) < 0$ $Ax = b$ is infeasible

3° suppose $f_i(x) < 0$ $Ax = b$ is feasible
$g(\lambda, \nu) = \inf\{\sum \lambda_i f_i(x) + \nu^T(Ax - b)\}$
< 0 when $\lambda_i \ge 0$ $\lambda \ne 0$
$\therefore \lambda_i \ge 0$ $\lambda \ne 0$ $g(\lambda, \nu) \ge 0$ is infeasible

4° suppose $\lambda \succeq 0$ $\lambda \ne 0$ $g(\lambda, \nu) \ge 0$ is infeasible

$\lambda \succeq 0$ $\mathbf{1}^T\lambda = 1$ $g(\lambda, \nu) \ge 0$ is infeasible
$\therefore \left\{\begin{array}{l}\max\ g(\lambda, \nu) \\ \text{s.t. } \lambda \succeq 0,\ \mathbf{1}^T\lambda = 1\end{array}\right\} < 0$
$\therefore p^* = d^* < 0$
$\therefore f_i(x) < 0$ $Ax = b$ is feasible

$\therefore \left\{\begin{array}{l} f_i(x) < 0 \quad Ax = b \\ \lambda \succeq 0 \quad \lambda \ne 0 \quad g(\lambda, \nu) \ge 0 \end{array}\right\}$ Strong alternative
exactly one of the two is feasible.
if $\exists x \in \text{relint } D$ with $Ax = b$

4. Strong alternatives for nonstrict inequalities

$f_i(x) \le 0 \quad Ax=b$

$\Downarrow$

$\min \ s$

s.t. $f_i(x) - s \le 0$

$Ax=b$

$$\inf_{x,s} L = \inf_{x,s}\left\{ s + \sum_i \lambda_i (f_i(x)-s) + v^T(Ax-b) \right\}$$

$$= \begin{cases} g(\lambda,v) & \text{if } 1^T\lambda = 1 \\ -\infty & \text{otherwise} \end{cases}$$

$\Downarrow$

$\max. \quad g(\lambda,v)$

s.t. $\lambda \ge 0 \quad 1^T\lambda = 1$

$g(\lambda^*,v^*) = p^* \quad \lambda^* \ge 0 \quad 1^T\lambda^* = 1$

suppose $\exists\ \tilde{x} \in \operatorname{relint} D$ with $A\tilde{x}=b$

choose $s > \max_i f_i(x)$ is strictly feasible

$\therefore\ p^* = d^*$

1° suppose the inequality system is infeasible

$p^* = d^* > 0$

(λ^*, v^*) satisfies $\lambda \ge 0 \quad g(\lambda,v) > 0$

2° suppose the inequality system is feasible

$p^* = d^* \le 0$

$g(\lambda,v) = \inf_x \left(\sum \lambda_i f_i(x) + v^T(Ax-b) \right) > 0$ is infeasible for $\lambda \ge 0$

3° suppose $\lambda \ge 0 \quad g(\lambda,v) > 0$ is infeasible

$\lambda \ge 0 \quad g(\lambda,v) > 0 \quad 1^T\lambda = 1$ is infeasible

$\therefore$ for $\lambda \ge 0 \quad 1^T\lambda = 1 \qquad d^* = p^* = g(\lambda^*, v^*) \le 0$

$\therefore$ the original problem is feasible

4° suppose $\lambda \ge 0 \quad g(\lambda,v) > 0$ is feasible

$p^* = d^* = \max_{\lambda,v}\{ g(\lambda,v),\ \lambda \ge 0\ \ 1^T\lambda = 1 \} > 0 \qquad g(\lambda,v)$ homogeneous in λ, v

$\therefore$ the original problem is infeasible

$$\left\{ \begin{array}{l} \lambda \ge 0 \quad g(\lambda,v) > 0 \\ f_i(x) \le 0 \quad Ax=b \end{array} \right\}$$

Strong alternatives

exactly one of these is feasible

if $\exists\ x \in \operatorname{relint} D$ s.t. $Ax=b$

- Generalized inequalities

$$\begin{aligned} \min.\ & f_0(x) \\ \text{s.t.}\ & f_i(x) \preceq_{K_i} 0 \\ & h_i(x) = 0 \end{aligned}$$

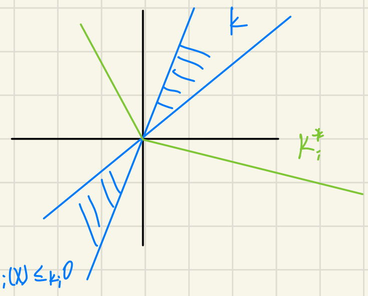
$\preceq_{K_i}$ is generalized inequality on $\mathbb{R}^{k_i}$

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

$$\mathcal{L}(x, \lambda, \cdots \lambda_m, v) = f_0(x) + \sum_i \lambda_i^T f_i(x) + v^T(Ax-b)$$ affine in λ, v

$$g(\lambda_1, \cdots, \lambda_m, v) = \inf_{x \in D} \mathcal{L}(x, \lambda, v)$$ point-wise infimum $\therefore$ concave

$$\begin{aligned} \max.\ & g(\lambda, v) \\ \text{s.t.}\ & \lambda_i \succeq_{K_i^*} 0 \end{aligned}$$ (so that $\lambda_i^T \cdot f_i(x) \le 0$ for feasible x)



lower bound property

$$\begin{aligned} g(\lambda, v) &= \inf_{x \in D} \left[f_0(x) + \sum_i \lambda_i^T f_i(x) + v^T(Ax-b) \right] \\ &\le f_0(x^*) + \sum_i \lambda_i^T f_i(x^*) + v^T(Ax^*-b) \\ &\le f_0(x^*) \quad \text{for } \lambda_i \succeq_{K_i^*} 0 \end{aligned}$$

Weak duality $p^* \le d^*$ always holds

Strong duality $p^* = d^*$ with constraint qualification (slater's condition e.g.)

eg. semi-definite programming

$$\begin{aligned} \min.\ & c^T x \\ \text{s.t.}\ & x_1 F_1 + \cdots + x_n F_n + G = 0 \end{aligned} \quad (F_i, G \in S^k)$$

Lagrangian multiplier is matrix $Z \in S^k$

$$\begin{aligned} \mathcal{L}(x, z) &= c^T x + tr\left(Z^T (x_1 F_1 + \cdots + x_n F_n + G)\right) \\ &= c^T x + tr\left(Z(x_1 F_1 + \cdots + x_n F_n + G)\right) \\ &= x_1 \cdot [c_1 + tr(ZF_1)] + \cdots + x_n [c_n + tr(ZF_n)] + tr(GZ) \end{aligned}$$

$L(x,z)$ is affine in x

$$g(z) = \inf_x = \begin{cases} tr(Gz) & tr(F_i z) + c_i = 0 \\ -\infty & \text{otherwise} \end{cases}$$

$$\Downarrow$$

max. $tr(Gz)$

s.t. $z \geq 0$ $\qquad (S^k)^* = S^k \qquad z \succeq_{S^k} 0$

$tr(F_i z) + c_i = 0$

$p^* = d^*$ if primal SDP is strictly feasible

- Complementary slackness for generalized inequality

assume $p^* = d^*$ attained at (x^*, λ^*, v^*)

$$f_0(x^*) = g(\lambda^*, v^*)$$
$$= \inf_x [f_0(x) + \sum_i \lambda_i^{*T} f_i(x) + \sum_i v_i \cdot h_i(x)]$$
$$\leq f_0(x^*) + \sum_i \lambda_i^{*T} f_i(x^*) + \sum v_i h_i(x^*)$$
$$\leq f_0(x^*)$$

$\therefore \lambda_i^{*T} f_i(x^*) = 0$

$\lambda_i^{*T} \succ_{k_i^*} 0 \Rightarrow f_i(x^*) = 0 \qquad f_i(x^*) \prec_{k_i^*} 0 \Rightarrow \lambda_i^* = 0$

- KKT condition for generalized inequality.

$f_i(x^*) \leq 0$

$h_i(x^*) = 0$

$\lambda_i^* \succeq_{k_i^*} 0$

$\lambda_i^{*T} f_i(x^*) = 0$

$\nabla f_0(x^*) + \sum_i Df_i(x^*) \lambda_i^* + \sum_i v_i^* \cdot \nabla h_i(x^*) = 0$