02/15/2018 Inequality constraints

$$\min f_0(x) \text{ s.t. } \begin{array}{l} f_i(x) \leq 0 \\ h_i(x) = 0 \end{array}$$

$$\mathcal{L}(x,\lambda,\nu) = f_0(x) + \sum \lambda_i f_i(x) + \sum \nu_i h_i(x)$$

$$g(\lambda,\nu) = \inf_x \mathcal{L}(x,\lambda,\nu) \qquad \text{(could be } -\infty\text{)}$$

would still like $g(\lambda,\nu) \leq f_0(x_{opt})$

$\leq \mathcal{L}(\tilde{x},\lambda,\nu)$ feasible $\tilde{x}$

$= f_0(\tilde{x}) + \sum \lambda_i \underbrace{f_i(\tilde{x})}_{\leq 0} + \sum \cancel{\nu_i h_i(\tilde{x})}$

$\leq f_0(\tilde{x})$ provided $\boxed{\lambda_i \geq 0}$ (denoted $\lambda \succeq 0$)

Dual problem: $\max g(\lambda,\nu)$ s.t. $\lambda \succeq 0$

Dual feasibility: (λ,ν) such that $\lambda \succeq 0$ and $g(\lambda,\nu) > -\infty$

<u>Def</u>. convex program: min a convex function over a convex set

→ the dual problem is <u>always</u> a convex program

Assume that strong duality holds.
Then

$$f_0(x_{opt}) = g(\lambda_{opt}, \nu_{opt}) = \inf_x \mathcal{L}(x, \lambda_{opt}, \nu_{opt})$$

$$\leq \mathcal{L}(x_{opt}, \lambda_{opt}, \nu_{opt})$$

$$= f_0(x_{opt}) + \sum (\lambda_{opt})_i f_i(x_{opt}) + \cancel{\sum (\nu_{opt})_i h_i(x_{opt})}$$

$$\Rightarrow \quad 0 \leq \sum \underbrace{(\lambda_{opt})_i}_{\geq 0} \underbrace{f_i(x_{opt})}_{\leq 0}$$

only possible if $\boxed{(\lambda_{opt})_i f_i(x_{opt}) = 0 \quad \forall i}$

so $f_i(x_{opt}) < 0 \Rightarrow (\lambda_{opt})_i = 0$
$(\lambda_{opt})_i > 0 \Rightarrow f_i(x_{opt})_i = 0$

Called complementary slackness.
At optimum, way to measure whether f_i is active or not.

KKT conditions: (f_0, f_i, h_i differentiable)

(1) $\nabla_x \mathcal{L}(x, \lambda, \nu) = 0 \Rightarrow \nabla f_0(x) + \sum \lambda_i \nabla f_i(x) + \sum \nu_i \nabla h_i(x) = 0.$
(2) $h_i(x) = 0$
(3) $f_i(x) \leq 0$
(4) $\lambda_i \geq 0$
(5) $\lambda_i f_i(x) = 0$

must hold at $(x_{opt}, \lambda_{opt}, \nu_{opt})$
when strong duality holds.

Prop. Assume $\begin{cases} f_0 & \text{diff, convex} \\ h_i & \text{affine} \\ f_i & \text{diff, convex} \end{cases}$

Then $(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$ obey KKT
$\Rightarrow$ they are optimal, and strong duality holds.

(Other conditions under which strong duality holds: "constraint qualification"
ex Slater's condition:
- some conditions on f_0, f_i, h_i
- KKT <u>not</u> assumed
- $\exists x$ s.t. $f_i(x) < 0$, $Ax = b$. (strict feasibility)

then strong duality holds)

ex. LP $\boxed{\begin{aligned} \min\ & c^T x \\ \text{s.t. } & Ax = b \\ & x \succeq 0 \end{aligned}}$ (standard form) $\rightarrow f_i(x) = -x_i$

$$\mathcal{L}(x, \lambda, \nu) = c^T x - \sum \lambda_i x_i + \nu^T (Ax - b)$$
$$= -b^T \nu + (c + A^T \nu - \lambda)^T x$$
$$g(\lambda, \nu) = \inf_x \mathcal{L}(x, \lambda, \nu)$$
$$= \begin{cases} -b^T \nu & \text{if } c + A^T \nu - \lambda = 0 \\ -\infty \end{cases}$$

Dual: $\max g(\lambda, \nu)$ s.t. $\lambda \succeq 0$

$\Leftrightarrow \max -b^T \nu$ s.t. $\begin{cases} A^T \nu - \lambda + c = 0 \\ \lambda \succeq 0 \end{cases}$

$\Leftrightarrow \boxed{\max -b^T \nu \text{ s.t. } A^T \nu + c \succeq 0}$

$\rightarrow$ another LP (inequality form)

Sensitivity analysis via duality.

Perturbed problem: $\min f_0(x)$
s.t. $f_i(x) \le u_i$
$h_i(x) = v_i$

with optimal value $f_0(x_{opt}(u,v)) = p(u,v)$

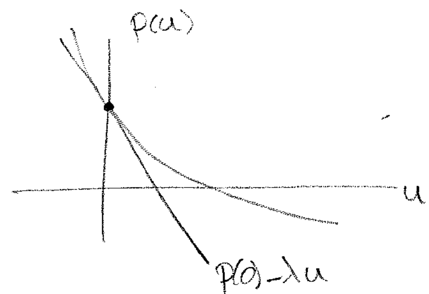
Assume feasible for small $|u_i|, |v_i|$
strong duality holds.

Prop. $p(u,v) \geq p(0,0) - (\lambda_{opt})^T u - (\nu_{opt})^T v$

Pf. $p(0,0) = g(\lambda_{opt}, \nu_{opt}) \overset{\text{feasible } x \text{ for the perturbed pb.}}{\leq} f_0(x) + \sum (\lambda_{opt})_i f_i(x) + \sum (\nu_{opt})_i h_i(x)$

$\leq f_0(x) + \sum (\lambda_{opt})_i u_i + \sum (\nu_{opt})_i v_i$

min over x, $\quad p(0,0) \leq p(u,v) + \lambda_{opt}^T u + \nu_{opt}^T v$



Prop. $(\lambda_{opt})_i = -\frac{\partial p}{\partial u_i}(0,0) \qquad (\nu_{opt})_i = -\frac{\partial p}{\partial v_i}(0,0)$

Pf Let $u_i = te_i$: $\quad p(te_i, 0) - p(0,0) \geq -(\lambda_{opt})_i t$

$t > 0$: $\quad \frac{p(te_i,0) - p(0,0)}{t} \geq -(\lambda_{opt})_i$

$t < 0$: $\quad \frac{p(te_i,0) - p(0,0)}{t} \leq -(\lambda_{opt})_i$

$\Rightarrow \frac{\partial p}{\partial u_i}(0,0) = -(\lambda_{opt})_i$ and similarly for $\partial p / \partial v_i$ $\square$

Cond: at $(x_{opt}, \lambda_{opt}, \nu_{opt})$:
recall $\underbrace{(\lambda_{opt})_i}_{\geq 0} \underbrace{f_i(x_{opt})}_{\leq 0} = 0$..

- if $f_i(x_{opt}) < 0$ (not saturated/inactive)
  then $(\lambda_{opt})_i = 0$ and perturbing f_i will not change the optimal objective value p.
- if $f_i(x_{opt}) = 0$ (saturated / active)
  then $(\lambda_{opt})_i \neq 0$ and $\frac{\partial p}{\partial u_i} = -(\lambda_{opt})_i$

(the larger $(\lambda_{opt})_i$, the more sensitive the objective to its constraint).

ex. (ULS): $\min \frac{1}{2} x^T x \; : \; Ax = b$

$$\mathcal{L}(x,\nu) = \frac{1}{2} x^T x + \nu^T (Ax - b)$$

$$\min \rightarrow \nabla_x \mathcal{L}(x,\nu) = x^T + \nu^T A = 0 \Rightarrow x = -A^T \nu$$

$$g(\nu) = \mathcal{L}(-A^T\nu, \nu) = \frac{1}{2} \nu^T A A^T \nu - \nu^T A A^T \nu - b^T \nu.$$
$$= -\frac{1}{2} \nu^T A A^T \nu - b^T \nu \quad \text{concave}$$

$$\max \rightarrow \nabla g(\nu) = -\nu^T A A^T - b^T = 0 \Rightarrow \nu_{opt} = -(AA^T)^{-1} b$$
$$\Rightarrow x_{opt} = (AA^T)^{-1} A^T b$$

and $(\nu_{opt})_i$ is the sensitivity of $a_i^T x = b_i$