Lecture 5 Strong duality

- Recap of last lecture
- strong duality
- complementary slackness conditions and KKT conditions

Primal optimization problem

domain of the optimization problem D

$$D = \operatorname{dom} f_0 \cap \bigcap_{i=1}^{m} \operatorname{dom} f_i \cap \bigcap_{j=1}^{p} \operatorname{dom} h_j$$

$$\inf f_0(x)$$

$$\text{s.t.} \quad f_i(x) \leq 0 \quad \forall i \in [m]$$

$$h_j(x) = 0 \quad \forall j \in [p]$$

feasible set

$p^* = -\infty$ if the problem is unbounded below

$p^* = +\infty$ if the feasible set is empty

$p^* \in \mathbb{R}$ if "everything" is fine

Lagrangian

$$L(x,\lambda,\mu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{n} \mu_i h_i(x)$$

Lagrangian dual function

$$g(\lambda,\mu) = \inf_x L(x,\lambda,\mu)$$

lemmas and observations

$\lambda \geq 0 \rightsquigarrow g(\lambda,\mu) \leq p^*$

dual program

$\max g(\lambda,\mu)$

$\lambda \geq 0$

best lower bound d^*

$d^* \leq p^* \rightsquigarrow$ weak duality

$\to$ always a convex program (no matter the primal)

Strong duality $d^* = p^*$

When does it hold? $\to$ LP, always

$\to$ general convex programs $\rightsquigarrow$ not always

Slater's condition = sufficient condition for strong duality to happen in case relint$(S) := \{x \in S : \exists \varepsilon > 0 \; B(\varepsilon, x) \cap \mathrm{aff}(S) \subseteq S\}$

Slater's condition

$\exists\, x \in \mathrm{relint}\, D$:

// feasible point in the relative interior that satisfies non-affine inequality strictly

$h_i(x) = 0$

$f_i(x) \leq 0$ (if f_i was affine)

$f_i(x) < 0$ (if f_i is not affine)

Theorem:

For convex programs

If Slater's condition hold then strong duality holds. (observation: note that in LP Slater's condition ⇝ program is feasible?)

We will prove a weaker version of Slater's condition, that is

$\exists\, \tilde{x} \in \mathrm{int}\, D$ s.t.

major assumption (for LP it does not hold)

① $f_i(x) < 0$ (for all inequality constraints including the affine ones)

minor assumption

② the equality constraints are linearly independent.

proof

consider the set $A = \{(t, v, u) \mid \exists x \text{ s.t. } t \geq f_0(x),\ v_i = h_i(x),\ u_i \geq f_i(x)\}$

observation 1

A is a convex set (easy to check since f_i are convex)

observation 2

if $(t, 0, 0) \in A$ then there is a feasible solution with value t.

observation 3

let p^* be the primal optimal value

then $\forall$ point $(s, 0, 0)$ with $s < p^*$

we know that $(s, 0, 0) \notin A$

let $B = \{(s, 0, 0) \mid s < p^*\}$

B is convex and by A and

B are disjoint $\rightsquigarrow$ separating hyperplane theorem !!!

$\exists\ \lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^n, \alpha \in \mathbb{R}$ s.t

$$\inf_{(t,v,u)\in A} \lambda^T u + \mu^T v + \alpha t \geq \sup_{(s,0,0)\in B} \alpha \cdot s$$

since $B = \{(s,0,0), s < p^*\}$

$$\sup_{(s,0,0)\in B} \alpha \cdot s = p^* \cdot \alpha$$

thus $\inf_{(t,v,u)\in A} \lambda^T u + \mu^T v + \alpha t \geq \alpha p^*$

since t and u are unbounded above for the infimum not to be $-\infty$ we need

$\lambda \geq 0$ and $\alpha \geq 0$ (we will prove it using Slater's condition)

now assume let $\alpha > 0$

then $\inf_{(t,v,u)} \left(\frac{\lambda}{\alpha}\right)^T u + \left(\frac{\mu}{\alpha}\right)^T v + t \geq p^*$

and because now all the multipliers are non-negative

$$\inf_{(t,v,u)} \left(\frac{\lambda}{a}^T\right) u + \left(\frac{\mu}{a}^T\right) v + t \geq p^* =$$

A is the "epigraph" of the optimization problem

$$\inf_x \sum_{i=1}^{m} \frac{\lambda_i}{a} f_i(x) + \sum_{i=1}^{n} \frac{\mu_i}{a} (A_i x - b_i) + f_0(x) \geq p^*$$

$$\Rightarrow \inf_x L\left(x, \frac{\lambda}{a}, \frac{\mu}{a}\right) \geq p^* \Rightarrow$$

$$\Rightarrow g\left(\frac{\lambda}{a}, \frac{\mu}{a}\right) \geq p^* \xRightarrow{\text{by weak duality}}$$

$$g\left(\frac{\lambda}{a}, \frac{\mu}{a}\right) = p^*$$

Now let's prove that Slater's condition ensures that $a > 0$.

proof (by contradiction)

assume that $a = 0$, by the definition of A we get

$$\sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{n} \mu_j (A_j^T x - b_i) \geq 0 \quad \forall x \in D$$

By Slater's condition there is a point that satisfies all non-affine inequalities strictly ① $\Rightarrow$

$$\sum_{i=1}^{m} \lambda_i \underbrace{f_i(\tilde{x})}_{\leq 0} + \sum_{j=1}^{n} \mu_j \overbrace{(A_j^T \tilde{x} - b_i)}^{0} \geq 0$$

$$\Rightarrow \boxed{\lambda_i = 0}$$

$$\Rightarrow \tilde{x} \in \mathrm{int}(B) \rightsquigarrow \tilde{x} + \varepsilon \cdot v$$

any direction is possible for sufficiently small ε

$$A = \begin{bmatrix} A_1^T \\ A_2^T \\ \vdots \\ A_j^T \end{bmatrix}$$

② rows of A are linearly independent $\Big\}$ $\sum_{j=1}^{n} \mu_j A_j^T \neq 0$ (because

$\mu \neq 0$

$(0, \mu, 0)$ should be a separating hyperplane.)

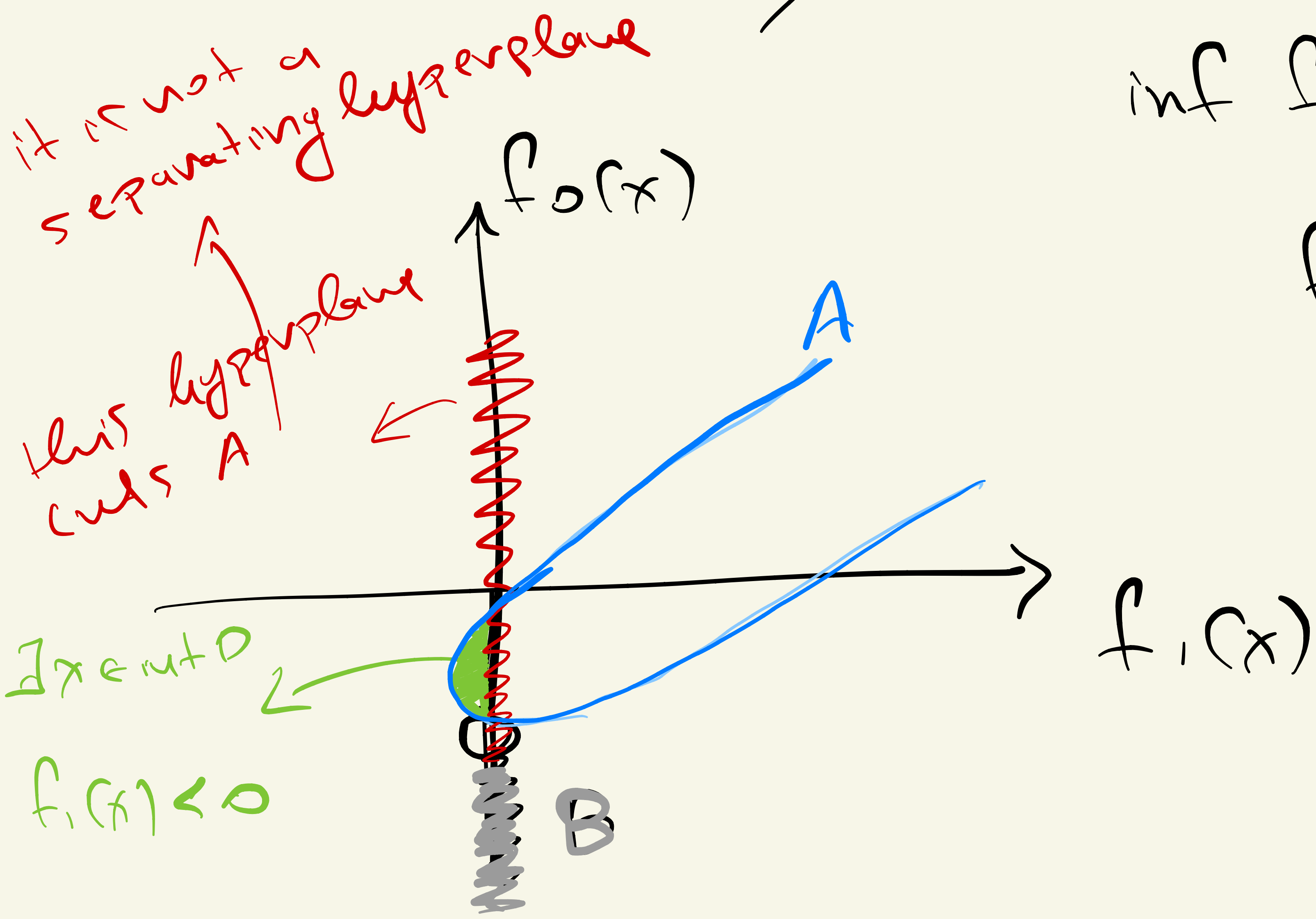
because I can go in any direction I want.

I can find x' s.t $\mu^T \cdot A x' < \underbrace{\mu^T \cdot A \tilde{x}}_{= b}$

$$\Rightarrow \sum_{j=1}^{n} \left(\mu^T \cdot A \cdot x' - b\right)_j < 0$$

$\Rightarrow$ contradiction (because $(0, \mu, 0)$ should separate A and B)

geometric intuition $\Longrightarrow$



$\inf f_0(x)$

$f_1(x) \leq 0$

Complementary slackness conditions

x^* is primal optimal

λ^*, μ^* are dual optimal

then: $f_0(x^*) = g(\lambda^*, \mu^*) =$

$$= \inf_x \mathcal{L}(x, \lambda^*, \mu^*) \leq$$

$$\leq \mathcal{L}(x^*, \lambda^*, \mu^*) =$$

$$= f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \underbrace{f_i(x^*)}_{\leq 0} + \sum_{j=1}^{n} \mu_j^* \underbrace{h_j(x^*)}_{= 0} \leq$$

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

the inequalities hold as equalities

$\overset{1st}{\Rightarrow} \inf_x \mathcal{L}(x, \lambda^*, \mu^*) = \mathcal{L}(x^*, \lambda^*, \mu^*)$

$\overset{2nd}{\Rightarrow} \lambda_i^* f_i(x^*) = 0 \ \forall i \in [m]$ complementary slackness

necessary conditions for a pair of primal and dual solutions to be optimal.

KKT conditions (when everything is differentiable and nice)

① $x^* = \arg\min \mathcal{L}(x, \lambda^*, \mu^*)$

$\Rightarrow \nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*) = 0 \Rightarrow$

$\Rightarrow \nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + \sum_{j=1}^{n} \mu_j \nabla h_j(x^*) = 0$

necessary conditions for any pair of primal and dual to be optimal.

primal feasibility

$f_i(x^*) \leq 0 \ \forall i \in [m] \quad h_j(x^*) = 0 \ \forall j \in [n]$

dual feasibility

$\lambda_i^* \geq 0 \ \forall i \in [m]$

complementary slackness

$\lambda_i^* \cdot f_i(x^*) = 0 \ \forall i \in [m]$

Lagrangian optimality

$\nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + \sum_{j=1}^{n} \mu_j \nabla h_j(x^*) = 0$

Lemma

If the problem is convex then the KKT conditions are also sufficient.

proof

$$g(\lambda^*, \mu^*) = \inf_x L(x, \lambda^*, \mu^*) =$$

$$= L(x^*, \lambda^*, \mu^*) =$$

$$= f_0(x^*) + \sum_{i=1}^{m} \lambda_i f_i(x^*) + \sum_{j=1}^{n} \mu_j h_j(x^*) =$$

$$= f_0(x^*)$$ ∎

for fixed λ^*, μ^*
$L(x, \lambda^*, \mu^*)$ is convex in x (since we assumed that the problem is convex)
$\Rightarrow$ local optimality $\Rightarrow$ global optimality

$\min \; e^{-x}$ ← convex

$\text{s.t} \; x^2/y \le 0$ → convex (quadratic over linear)

$D = \{(x,y) \text{ s.t } y > 0\}$

optimal value ⇝ $x = 0$ ⇝ $e^{-0} = 1 = p^*$

$$g(\lambda) = \inf_{\substack{(x,y)\\ y>0}} \left\{ e^{-x} + \lambda \cdot x^2/y \right\} = \begin{cases} -\infty & \lambda < 0 \\ 0 \;\; (x \to +\infty) & \lambda \ge 0 \end{cases}$$

$$\left.\begin{aligned} &\max \; g(\lambda) \\ &\text{s.t} \;\; \lambda \ge 0 \end{aligned} \;=\; \begin{aligned} &\max \; 0 \\ &\text{s.t} \;\; \lambda \ge 0 \end{aligned}\right\} \; d^* = 0$$

duality gap $p^* - d^* = 1 - 0 = 1$ ⇒

⇒ Slater condition should not be satisfied

indeed $\nexists (x,y) \in D \quad x^2/y < 0$