LECTURE 10

STRONG DUALITY

$$p_* := \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x)$$
$$c_j(x) = 0 \quad j \in E$$
$$c_j(x) \geq 0 \quad j \in I$$

see previous lecture

PRIMAL PROBLEM

$$d_* := \underset{\mu \in \mathbb{R}^m,\, \lambda \in \mathbb{R}^p}{\text{maximize}} \quad g(\mu, \lambda)$$
$$\lambda \geq 0$$

see previous lecture

DUAL PROBLEM

where

$$g(\mu, \lambda) := \inf_{x \in D} \quad L(x; \mu, \lambda)$$

and

$$L(x; \mu, \lambda) := f(x) - \sum_{j \in E} \mu_j c_j(x) - \sum_{j \in I} \lambda_j c_j(x)$$

$$D := \left\{ \bigcap_{j \in I \cup E} \operatorname{dom} c_j(x) \right\} \cap \operatorname{dom} f(x)$$

We always have the weak duality

$$p_* \geq d_*$$

Strong duality is said to hold if

$$p_* = d_*$$

Convex optimization problems

(Convex programs)

$$\begin{aligned} &\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x) \\ &c_j(x) \geq 0 \quad j \in I \\ &Ax = b \end{aligned}$$

where

f is convex

$-c_j$ is convex $\forall j \in I$

Remark

Above defn is equivalent to $f(x)$ is convex, and the feasible region F is a convex set.

$$F = \{x \mid c_j(x) \geq 0 \text{ and } Ax = b\}$$

In particular the dual problem.

$$\underset{\substack{\mu \in \mathbb{R}^m, \lambda \in \mathbb{R}^p \\ \lambda \geq 0}}{\text{maximize}} \quad g(\mu, \lambda) \quad \left(\equiv - \underset{\substack{\mu, \lambda \\ \lambda \geq 0}}{\text{minimize}} \; -g(\mu, \lambda) \right)$$

is convex, since $-g(\mu, \lambda)$ is convex.

Geometric View for Strong / Weak Duality

$$G := \{ (c_{i_1}(x), \cdot, c_{i_m}(x), c_{k_1}(x), \cdot, c_{k_p}(x), f(x)) \mid x \in D \}$$

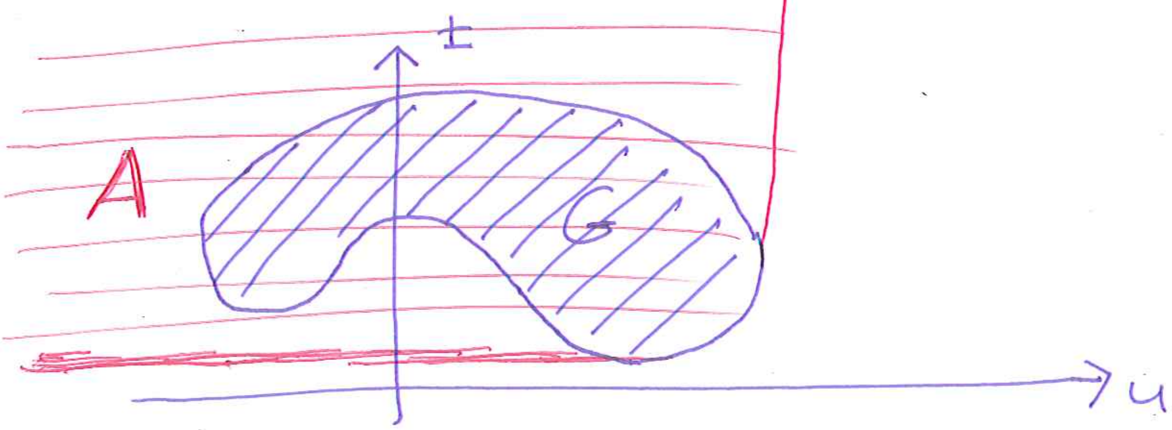
Above,

$$E = \{i_1, \dots, i_m\} \text{ and } I = \{k_1, \dots, k_p\}$$

$$A := \{ (v,u,t) \mid v \in \mathbb{R}^m, u \in \mathbb{R}^p, t \in \mathbb{R} \text{ such that } \exists x \in D \;\; f(x) \leq t, \;\; c_{i_l}(x) = v_l \;\; l = 1, \dots, m, \;\; c_{k_l}(x) \geq u_l \;\; l = 1, \dots, p \}$$

Illustration (one inequality case)

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \; f(x)$$
$$c(x) \geq 0$$



Primal problem

$$P_* = \inf \{ t \mid (0,0,t) \in A \}$$
$$= \inf \{ t \mid \exists (v,u) \text{ s.t. } (v,u,t) \in G, \; u \geq 0, \; v = 0 \}$$

Lagrange dual function

$$g(\mu,\lambda) = \inf\left\{ \overbrace{(-\mu,-\lambda,1)^T(v,u,t)}^{-\mu^T v - \lambda^T u + t} \;\middle|\; (v,u,t) \in G \right\}$$

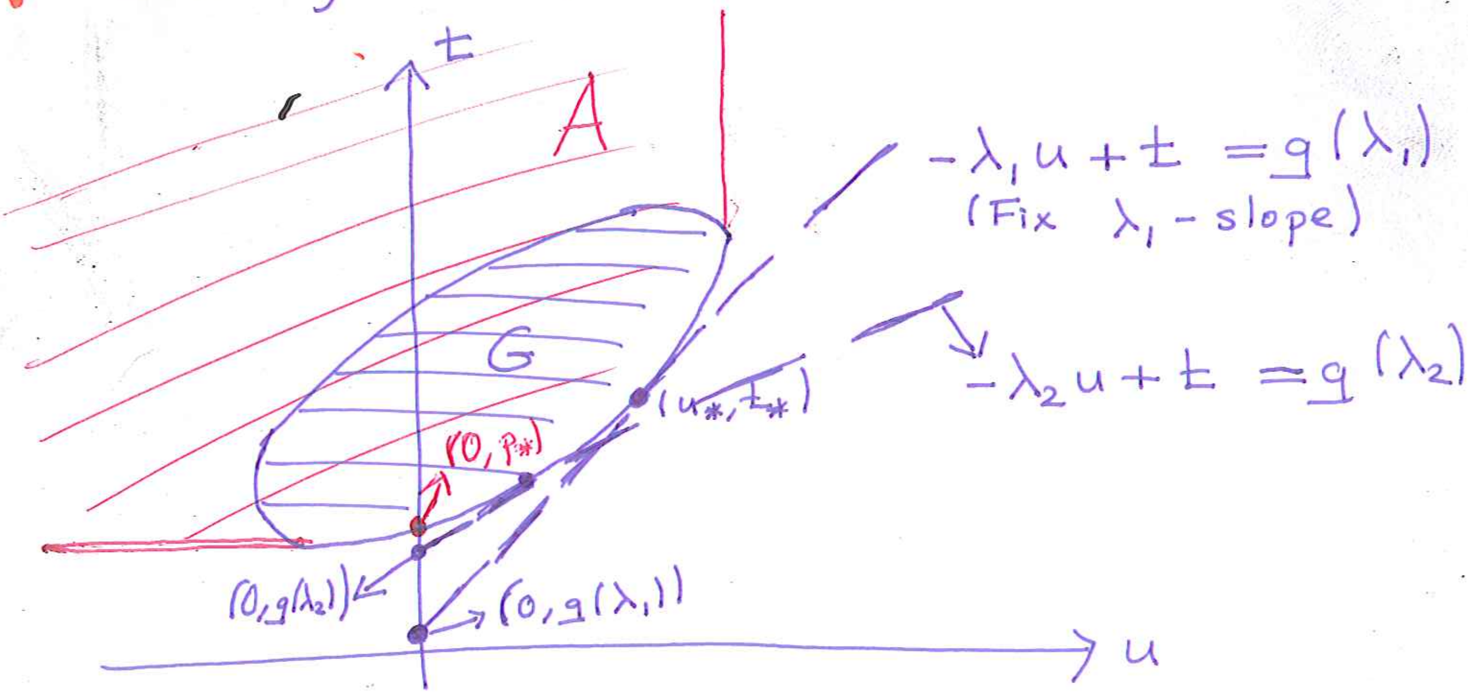
If infimum is attained

(i) $\overbrace{(-\mu,-\lambda,1)^T(v,u,t)}^{-\mu^T v - \lambda^T u + t} \geq g(\mu,\lambda) \quad \forall (v,u,t) \in G$

(ii) $\underbrace{(-\mu,-\lambda,1)^T(v_*,u_*,t_*)}_{-\mu^T v_* - \lambda^T u_* + t_*} = g(\mu,\lambda) \quad \exists (v_*,u_*,t_*) \in G$

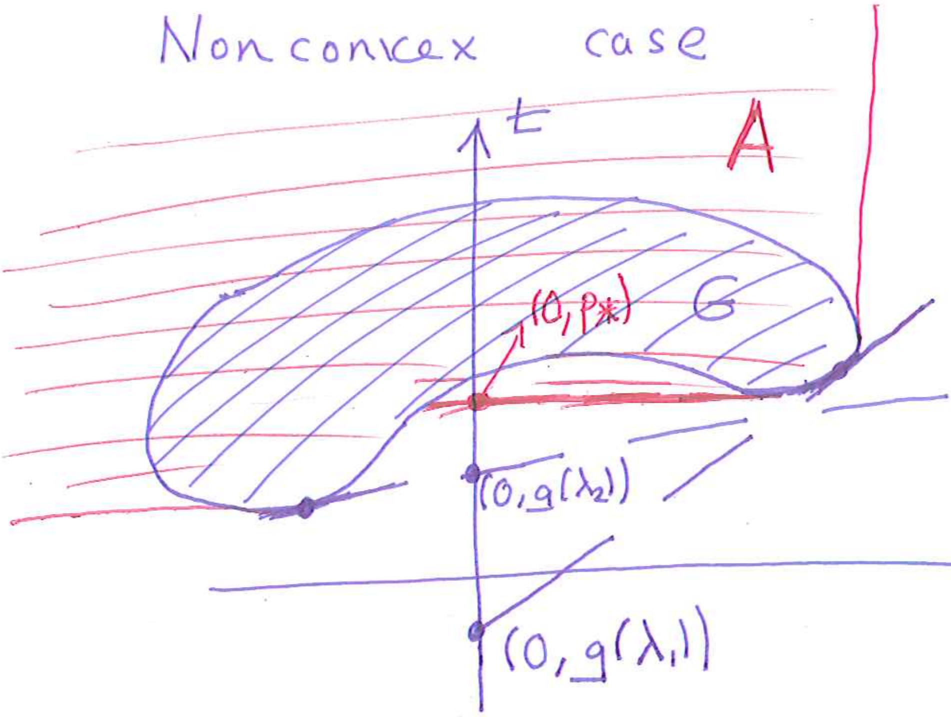
Illustration

Convex case (A has to be convex, but for simplicity suppose G is also)



$$\left[d_* = \sup_{\lambda \geq 0} g(\lambda) \right] = p_*$$

Nonconvex case



$-\lambda_1 u + t = g(\lambda_1)$

$-\lambda_2 u + t = g(\lambda_2)$

$$\left[d_* = \sup_{\lambda \geq 0} g(\lambda) \right] \neq p_*$$

Suppose the infimum

$$p_* = \inf \{ t \mid (0,0,t) \in A \}$$

is attained.

For each $(\mu, \lambda) \in \operatorname{dom} g$ s.t. $\lambda \geq 0$

$$p_* = (-\mu, -\lambda, 1)^T \cdot (0, 0, p_*) \geq g(\mu, \lambda)$$

(intuitively $-\lambda^T u - \mu^T v + t = g(\mu,\lambda)$ tangent to G at $(0,0,p)$

Strong duality holds $\iff$ Above holds with equality $\exists (\mu,\lambda) \in \operatorname{dom} g$ s.t. $\lambda \geq 0$.

Interior

Let D be a given set.

$x_* \in \text{int}(D)$ if $\exists \delta > 0$ such that

$\cancel{x_* \in}$ $B(x_*, \delta) \subseteq D$.

Relative interior

$x_* \in \text{relint}(D)$ if $\exists \delta > 0$ such that

$B(x_*, \delta) \cap \text{Aff}(D) \subseteq D$.

$$\text{Aff}(D) := \{ \theta_1 x_1 + \cdots + \theta_K x_K \mid K \in \mathbb{Z}^+,\ x_1, \ldots, x_K \in D,\ \theta_1, \ldots, \theta_K \in \mathbb{R} \text{ s.t. } \sum_{j=1}^{K} \theta_j = 1 \}$$

e.g. $\text{Aff}\left(\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \right) = \{ t(1,0) + (1-t)(0,1) \mid t \in \mathbb{R} \}$

$= \{ (0,1) + t(1,-1) \mid t \in \mathbb{R} \}$

line through $(1,0)$ and $(0,1)$

Ex

$D := \{ (x,0) \mid x \in [0,1] \}$

$\text{int } D = \phi$

$\text{Aff } D = \{ (x,0) \mid x \in \mathbb{R} \}$

$\text{relint } D = \{ (x,0) \mid x \in (0,1) \}$

THM (Slater Condition)

Suppose that the primal problem is convex, and that there exists $x_* \in \text{Relint}(D)$ such that

$$(i)\ Ax_* = b \quad \text{and} \quad (ii)\ c_j(x_*) > 0 \ \forall j \in I.$$

Then $p_* = d_*$.

Remark

If $\exists S \subseteq I$ such that for each $j \in S$

$$c_j(x) = a_j^T x + b_j$$

for some $a_j, b_j \in \mathbb{R}$, then condition (ii) above can be replaced by

$$c_j(x_*) > 0 \ \forall j \in I \setminus S \quad \text{and} \quad c_j(x_*) \geq 0 \ \forall j \in S.$$

Ex

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \frac{1}{2} x^T H x + g^T x$$

$$Ax = b$$

$$x \geq 0$$

H - $n \times n$ symmetric PD

A - $m \times n$

$g \in \mathbb{R}^n$, $b \in \mathbb{R}^m$

$$L(x;\mu,\lambda) = \frac{1}{2}x^T H x + \underline{g}^T x - \mu^T(Ax-b) - \lambda^T x$$

$$\underline{g}(\mu,\lambda) = \inf_{x\in\mathbb{R}^n} \frac{1}{2}x^T H x + \underline{g}^T x - \mu^T(Ax-b) - \lambda^T x$$

$$= \mu^T b + \inf_{x\in\mathbb{R}^n} \frac{1}{2}x^T H x + (\underline{g} - A^T\mu - \lambda)^T x$$

Optimal x_* for the minimization problem

$$x_* = H^{-1}(A^T\mu + \lambda - \underline{g})$$

so

$$\underline{g}(\mu,\lambda) = \mu^T b - \frac{1}{2}(A^T\mu + \lambda - \underline{g})\, H^{-1}(A^T\mu + \lambda - \underline{g})$$

Dual problem

$$\begin{array}{ll} \underset{\mu\in\mathbb{R}^m,\ \lambda\in\mathbb{R}^n}{\text{maximize}} & \mu^T b - \frac{1}{2}(A^T\mu+\lambda-\underline{g})H^{-1}(A^T\mu+\lambda-\underline{g}) \\ & \lambda \geq 0 \end{array}$$

If there exists an x such that

$$Ax = b \quad \text{and} \quad x \geq 0$$

the primal and dual problems have the same optimal value.

Ex (Matrix games)

P1 makes a choice from $\{1, 2, \ldots, n\}$

P2 makes a choice from $\{1, 2, \ldots, m\}$

P - $n \times m$ matrix, p_{kl} corresponds to the payment P1 makes to P2 when P1 chooses k and P2 chooses l.

u, v — (discrete) probability distribution associated with choices of P1, P2

$$\sum_{i=1}^{n} u_i = \sum_{i=1}^{m} v_i = 1 \quad \text{and} \quad u, v \geq 0.$$

Expected payment

$$\sum_{k=1}^{n} \sum_{l=1}^{m} u_k v_l p_{kl} = u^T P v$$

P2 aims to maximize by choosing v

P1 aims to minimize by choosing u

Suppose P2 plays the best possible (given u-play of P1)

$$\underset{v}{\text{maximize}} \; u^T P v = \max_{j=1,\ldots,m} (P^T u)_j$$

The best P1 can do

$$\underset{u}{\text{minimize}} \; \underset{v}{\text{maximize}} \; u^T P v$$

$$= \underset{u}{\text{minimize}} \left\{ \max_{j=1,\ldots,m} (P^T u)_j \right\}$$

that is

$$p_* := \underset{u \in \mathbb{R}^n, t \in \mathbb{R}}{\text{minimize}} \quad t$$

(P1) $e^T \cdot u = 1$
$u \geq 0$
$P^T u \leq t e$

Conversely, suppose P1 plays the best possible (given v - play of P_2)

$$\underset{u}{\text{minimize}} \; u^T P v = \min_{j=1,\ldots,n} (P v)_j$$

The best P2 can do

$$\underset{v}{\text{maximize}} \left\{ \min_{j=1,\dots,n} (Pv)_j \right\}$$

that is

$$d_* := \underset{v \in \mathbb{R}^m,\ t \in \mathbb{R}}{\text{maximize}} \quad t$$

P2
$$e^T \cdot v = 1$$
$$v \geq 0$$
$$P \cdot v \geq t \cdot e$$

P1 and P2 are primal dual problems (i.e., P2 is the dual of P1) - Verification exercise.

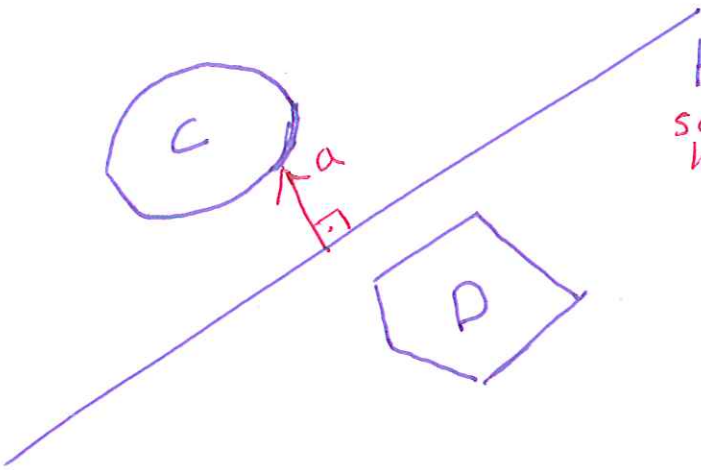
Slater conditions guarantee $p_* = d_*$.

LEMMA (Separating Hyperplane)

Let C and D be nonempty convex sets in $\mathbb{R}^n$ such that $C \cap D = \emptyset$. There exists a nonzero $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$ s.t.

(i) $a^T x \geq b \quad \forall x \in C$, and

(ii) $a^T x \leq b \quad \forall x \in D$



PROOF OF THM (Slater Conditions)

Assuming - Relint(D) = Int(D) and Rank(A) = m

Consider the set

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

Without loss of generality, we assume p_* is finite, i.e., $p_* \neq \infty$ is excluded since $\exists$ feasible points by assumption, if $p_* = -\infty$, then $d_* = -\infty$ by weak duality.

Notice that $A \cap B \neq \emptyset$: Assume $(v, u, t) \in A \cap B$, then $u = v = 0$ and $t < p_*$. Furthermore, since $(v, u, t) \in A$, we have $c_{i_\ell}(x) = 0 \quad \ell = 1, \ldots, m$
$c_{k_\ell}(x) \geq 0 \quad \ell = 1, \ldots, p$ and $f(x) \leq t < p_* \; \exists x \in D$, which is a contradiction!

Due to assumption $f(x)$ is convex, $c_{k\ell}(x)$ $\ell = 1, \ldots, p$ are convex, the set A is convex.
Thus by the lemma above $\exists (\lambda, \mu, n) \neq 0$ such that

$$(+) \quad (-\mu, -\lambda, n)^T (v, u, t) \geqslant \alpha \qquad \forall (v, u, t) \in A$$

$$(++) \quad (-\mu, -\lambda, n)^T (v, u, t) \leqslant \alpha \qquad \forall (v, u, t) \in B$$

Now $\lambda, n \geqslant 0$; because otherwise if $(v, u, t) \in A$, so does $(v, \tilde{u}, \tilde{t}) \in A$ for every $\tilde{u} \geqslant \tilde{u}$ and $\tilde{t} \leqslant \tilde{t}$ implying $-\mu^T \tilde{v} - \lambda^T \tilde{u} + n\tilde{t}$ is unbounded below over A contradicting $(+)$.

Moreover, due to $(++)$

$$n s \leqslant \alpha \qquad \forall s \text{ s.t. } s < p_*$$

$$\Longrightarrow n p_* \leqslant \alpha.$$

Now exploiting $(+)$, for each $x \in D$, we have

$$(+++) \quad -\sum_{j \in E} \mu_j \overset{Ax-b}{c_j(x)} - \sum_{j \in I} \lambda_j c_j(x) + nt \geqslant \alpha \geqslant n p_*$$

First assume $n \neq 0$. But then

$$L\left(x; \frac{\mu}{n}, \overset{\geqslant 0 \ (\text{i.e., } \lambda, n \geqslant 0)}{\frac{\lambda}{n}}\right) \geqslant p_*$$

for every $x \in D$ implying

$$g\left(\frac{\mu}{n}, \overset{\geqslant 0}{\frac{\lambda}{n}}\right) \geqslant p_* \Longrightarrow d_* \geqslant p_*$$

Combining this with weak duality yields $d_* = p_*$ as desired.

Now assume $\eta = 0$. There exists $\tilde{x} \in \mathrm{Int}(D)$ such that $c_j(\tilde{x}) > 0$ and $A\tilde{x} = b$ by assumption.

But then from (+++)

$$-\sum_{j \in E} \mu_j \overset{A\tilde{x}-b}{c_j(\tilde{x})} - \sum_{j \in I} \lambda_j c_j(\tilde{x}) \geqslant 0$$

$$\Longrightarrow -\sum_{j \in I} \lambda_j c_j(\tilde{x}) \geqslant 0$$

$$\Longrightarrow \lambda = 0.$$

Consequently, $\mu \neq 0$. From (+++) for each $x \in D$

$$(++++) \quad -\sum_{j \in E} \mu_j \overset{(A\tilde{x}-b)}{c_j(x)} \geqslant 0.$$

But $\tilde{x} \in \mathrm{Int}\, D$ such that $A\tilde{x} = b$.

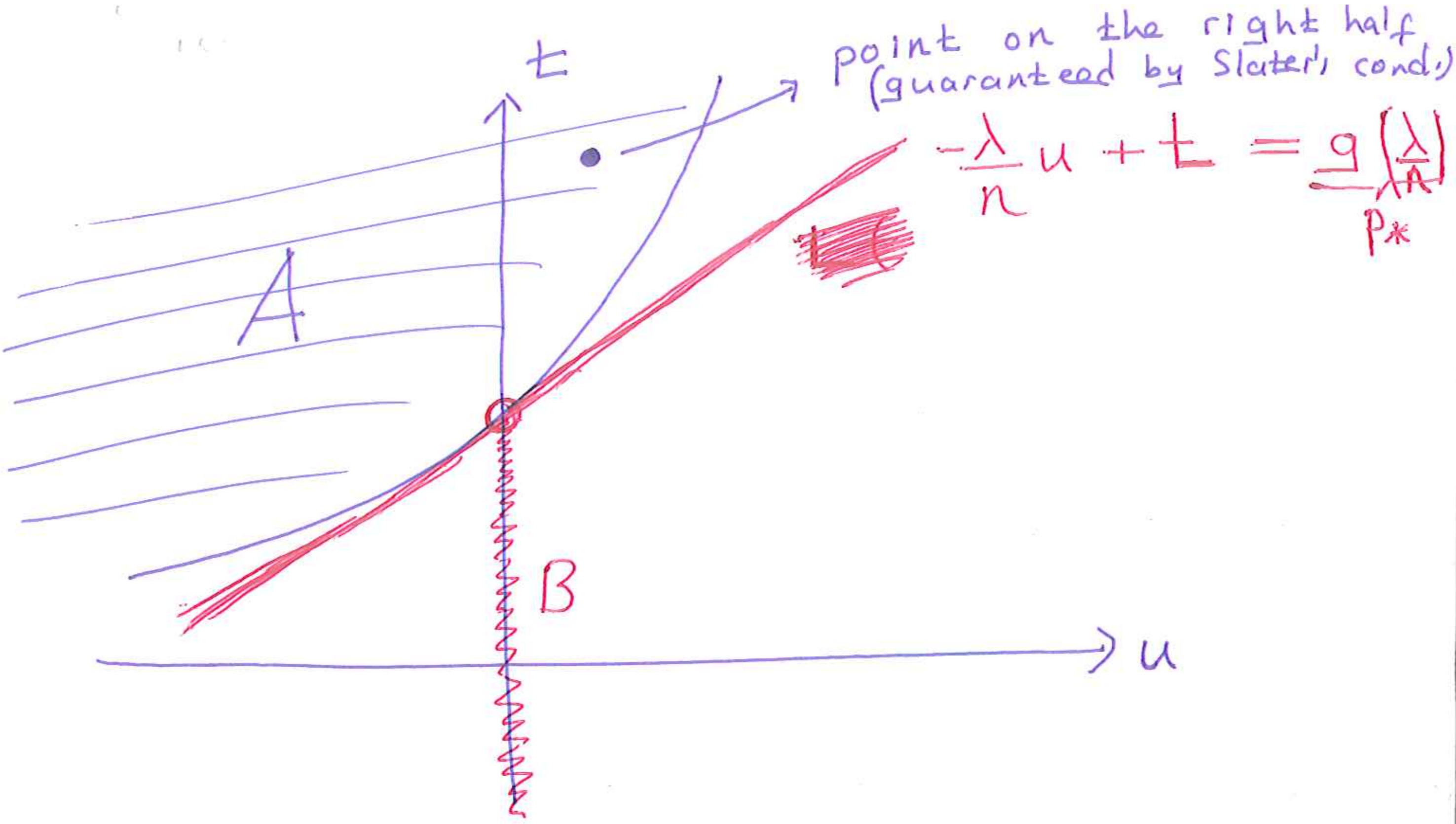
Due to assumption $\mathrm{Rank}(A) = m$, there exists p such that $Ap = \mu$. For each $\alpha > 0$ small enough $\tilde{x} + \alpha p \in \mathrm{Int}\, D$ and

$$-\underbrace{\mu^T\big(A(\tilde{x}+\alpha p) - b\big)}_{\sum_{j \in E} \mu_j c_j(\tilde{x}+\alpha p)} = -\alpha \mu^T A p$$

$$= -\alpha \|\mu\|^2 < 0.$$

This contradicts (++++). Thus $\eta \neq 0$.

$\square$


point on the right half
(guaranteed by Slater's cond.)
$-\frac{\lambda}{n}u + t = g\left(\frac{\lambda}{n}\right)$
p_*
t
A
B
u