- **Duality** (standard form)

Primal LP

maximize $c_1x_1 + c_2x_2 + \cdots + c_nx_n$

subject to
$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \le b_1$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \le b_2$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \le b_m$$
$$x_1, x_2, \cdots, x_n \ge 0.$$

Dual LP

minimize $b_1y_1 + \cdots + b_my_m$

subject to
$$a_{11}y_1 + a_{21}y_2 + \cdots + a_{m1}y_m \ge c_1$$
$$a_{12}y_1 + a_{22}y_2 + \cdots + a_{m2}y_m \ge c_2$$
$$\cdots$$
$$a_{1n}y_1 + a_{2n}y_2 + \cdots + a_{mn}y_m \ge c_n$$
$$y_1, y_2, \cdots, y_m \ge 0.$$

- **Matrix form.**

$m \times n$ matrix $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & \cdots & \cdots & a_{mn} \end{bmatrix}$, $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$, $\vec{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{R}^n$ $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{R}^m$. $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \in \mathbb{R}^m$.

A^T = transpose of A

$= \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & & \ddots & \vdots \\ a_{1n} & & \cdots & a_{mn} \end{bmatrix}$ $n \times m$ matrix

primal LP

$\vec{c}^T\vec{x} = \vec{c} \cdot \vec{x}$

maximize $\vec{c}^T\vec{x}$

subject to $A\vec{x} \le \vec{b}$

$\vec{x} \ge \vec{0}$.

dual LP.

minimize $\vec{b}^T\vec{y}$

subject to $A^T\vec{y} \ge \vec{c}$

$\vec{y} \ge \vec{0}$

Lecture 10 (TuTh)

Duality. Vanderbei sec 5.1 , 5.2, 5.3

- - Where does duality come from? Related to Vanderbei sec 5.9 and 5.10

Next lecture : - weak duality. Vanderbei 5.3

- Where does the duality come from?

To give a motivation for duality, we consider:

- Penalty method for constraint optimization problem.

f : function of $\vec{x} \in \mathbb{R}^n$ ...

$C \subset \mathbb{R}^n$ subset (constraint).

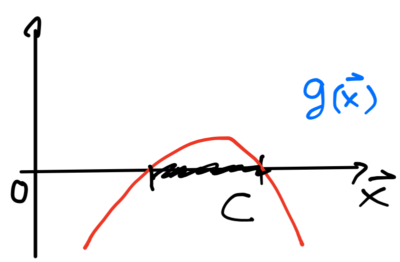
$$\left\{ \begin{array}{l} \text{maximize } f(\vec{x}) \\ \underbrace{\text{subject to } \vec{x} \in C}_{\text{hard constraint}} \end{array} \right\} \Leftarrow \text{maximize } f(\vec{x}) + \underbrace{p}_{\text{penalty}}(\vec{x})$$

where $p(\vec{x}) = \begin{cases} 0 & \text{for } \vec{x} \in C \\ -\infty & \text{for } \vec{x} \notin C \end{cases}$

Relaxation of penalty:

maximize $f(\vec{x}) + g(\vec{x})$

for the relaxed penalty



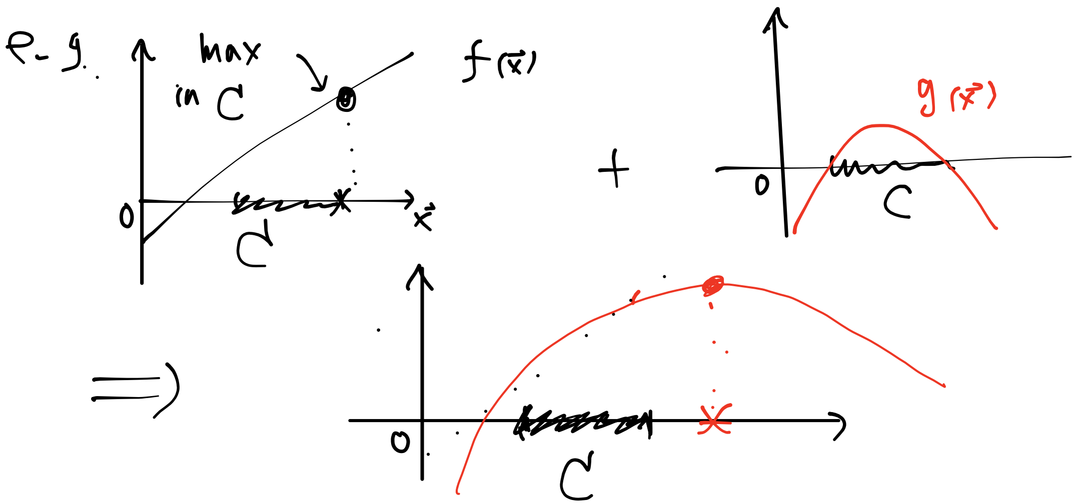
such that $g(\vec{x}) \geq 0$ if $\vec{x} \in C$

$g(\vec{x}) < 0$ if $\vec{x} \notin C$.

and $g(\vec{x})$ gets more negative

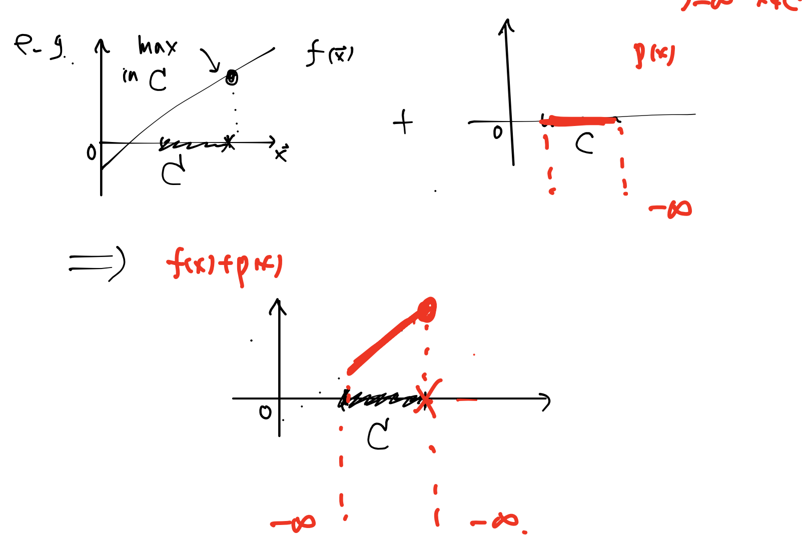
as $\vec{x}$ goes away from C.

- With soft penalty



- With hard penalty.
(infinite)

$p(x) = \begin{cases} 0 & x \in C \\ -\infty & x \notin C \end{cases}$



- In the relaxed penalization, it is ok not to satisfy the constraint, but still it is preferable to stay in or near the constraint.
- The optimal solution of the relaxed problem can still be close to the optimal solution of unrelaxed problem, if an appropriate relaxation is made.

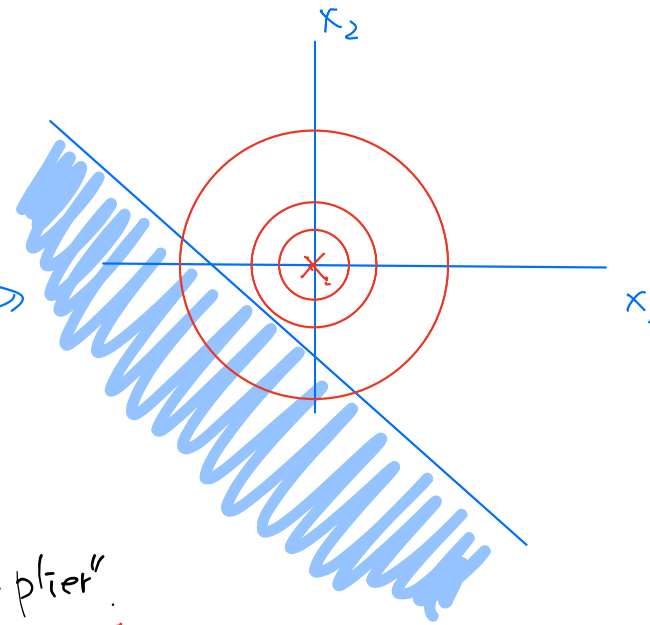
* $g(\vec{x})$ is like giving penalty on **how much** you **violate** the constraint.
If a policy maker sets up such a **penalty** "right",
then he/she gets the same effect as **enforcing** the constraint **strictly**.

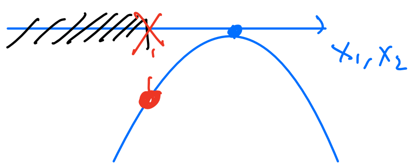
∘ A typical way to construct a penalty function.

Lagrange multiplier (adjusting the penalization)

EX
$$\boxed{\begin{array}{ll}\text{maximize} & -x_1^2-x_2^2\\ \text{subject to} & x_1+x_2\le -1\end{array}}$$

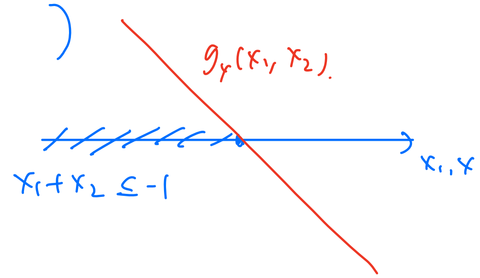
$f(x) = -x_1^2 - x_2^2$, $C = \{x_1+x_2 \le -1\}$.



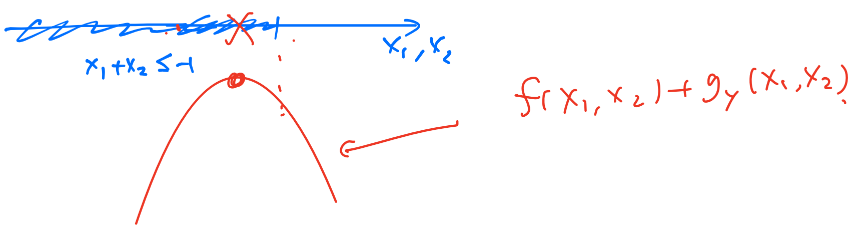


Consider, for fixed $y \geq 0$ "Lagrange multiplier".

$$g_y(x_1,x_2) = y\left(-1-(x_1+x_2)\right)$$



maximize $-x_1^2-x_2^2 + g_y(x_1,x_2)$



Remark

The Lagrange multiplier here is the same one as in the multivariable calculus;

At maximum point of $H(x)+yG(x)$,

we have

$$\nabla\left(H(x)+yG(x)\right)=0$$

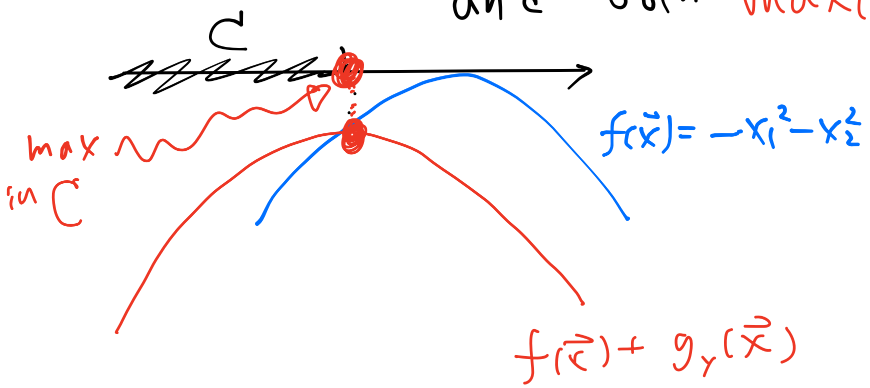
so $\nabla H(x) = -y\nabla G(x)$.

change in $y \Rightarrow$ change in $\max \left(f(\vec{x}) + g_y(\vec{x}) \right)$

At a "right" $\boxed{y}$ (in fact $y = 1$ in this example)

$$\max \left(f(\vec{x}) + g_1(\vec{x}) \right) = \max_{\vec{x} \in C} f(x).$$

and both maxima occur at the SAME $\vec{x}$.



WITH THE SAME OPTIMAL OBJECTIVE VALUE!

* y is like giving a "fee / price" rate on how much you violate the constraint.
If a policy maker sets up such a rate "right", then he/she gets the same effect as forcing the constraint.

- Penalty method gives motivation for duality in LP. In LP, the relaxed problem does not give us the same optimal solution as the original problem. BUT, it gives the same optimal objective value for a good choice of Lagrange multipliers.

"Lagrange multiplier for LP".

- Back to LP.

maximize $\vec{c}^T\vec{x}$

subject to $A\vec{x} \le \vec{b}$ ← m inequalities

$\vec{x} \ge \vec{0}$ ← n inequalities $x_1 \ge 0, x_2 \ge 0, \dots, x_n \ge 0$

$m+n$ constraints.

transpose of vector: $\vec{c}^T = [c_1, \cdots, c_n]$

$\vec{c}\cdot\vec{x} = [c_1, \cdots, c_n]\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \vec{c}^T\vec{x}$.

Consider

Lagrange multiplier $\vec{y} \in \mathbb{R}^m$ $\vec{y} \ge \vec{0}$ & $\vec{w} \in \mathbb{R}^n$ $\vec{w} \ge \vec{0}$

$(y_1, \dots, y_m)$ $(w_1, \dots, w_n)$

relaxed penalty. $g_{\vec{y},\vec{w}}(\vec{x}) = \boxed{\vec{y}^T}(\vec{b} - A\vec{x}) + \boxed{\vec{w}^T}\vec{x}$

$$\vec{y}^T(\vec{b} - A\vec{x})$$
$$= y_1\left(b_1 - \sum_{j=1}^n a_{1j}x_j\right) + y_2\left(b_2 - \sum_{j=1}^n a_{2j}x_j\right) + \cdots + y_m\left(b_m - \sum_{j=1}^n a_{mj}x_j\right)$$

$$\vec{w}^T\vec{x} = w_1x_1 + w_2x_2 + \cdots + w_nx_n$$

Note as $\vec{y}, \vec{w} \ge 0$, $g_{\vec{y},\vec{w}}(\vec{x}) \ge 0$ if $\vec{x} \ge \vec{0}$ & $A\vec{x} \le \vec{b}$

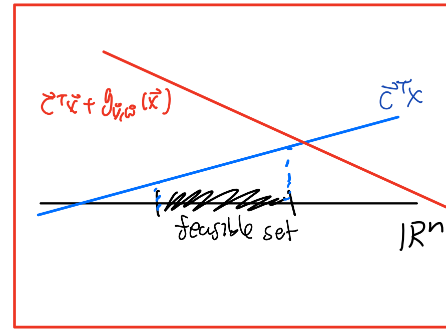
And the relaxed problem: for given $\vec{c}, A, \vec{b}, \vec{y}, \vec{w}$,

$$\underset{\vec{x}}{\text{maximize}}\left[\vec{c}^T\vec{x} + g_{\vec{y},\vec{w}}(\vec{x})\right].$$ (with no constraint).

Note relaxed problem has more options than the original constraint problem.

Recall $g_{\vec{y},\vec{w}}(\vec{x}) \geq 0$ if $\vec{x} \geq \vec{0}$ and $A\vec{x} \leq \vec{b}$

(because $\vec{y} \geq \vec{0}$, $\vec{w} \geq \vec{0}$)



Therefore. $\boxed{\vec{c}^T\vec{x} + g_{\vec{y},\vec{w}}(\vec{x}) \geq \vec{c}^T\vec{x}, \quad \text{if} \quad \vec{x} \geq \vec{0} \ \& \ A\vec{x} \leq \vec{b}.}$

Let $F(\vec{y},\vec{w}) = \max_{\vec{x}} \left[\vec{c}^T\vec{x} + g_{\vec{y},\vec{w}}(\vec{x})\right]$ the maximum value of the relaxed problem for fixed $\vec{y}, \vec{w} \geq \vec{0}$

Then,

$$F(\vec{y}, \vec{w}) \geq \max_{\substack{A\vec{x} \leq \vec{b} \\ \vec{x} \geq \vec{0}}} \vec{c}^T\vec{x}.$$

We used here two obvious facts.

① $\max_{\vec{x}} H(\vec{x}) \geq \max_{\vec{x} \in C} H(\vec{x})$

② If $H(\vec{x}) \geq G(\vec{x})$ for $x \in C$. then $\max_{\vec{x} \in C} H(\vec{x}) \geq \max_{\vec{x} \in C} G(\vec{x})$

Q How can we make the relaxed problem as close to the original problem as possible?

IDEA: To get the right Lagrange multiplier $\vec{y}, \vec{w}$,

$$\boxed{\begin{aligned} &\text{minimize} \quad F(\vec{y},\vec{w}) \\ &\text{subject to} \quad \vec{y}, \vec{w} \geq \vec{0}. \end{aligned}}$$

$$F(\vec{y},\vec{w}) = \max_{\vec{x}} \left(\vec{c}^T\vec{x} + \vec{y}^T(\vec{b} - A\vec{x}) + \vec{w}^T\vec{x}\right).$$

We will show that this problem is equivalent to the dual problem of the original LP.

Choose a correct statement:

A) For a given vector $\vec{a} \in \mathbb{R}^n$, it must be true that

$$\max_{\vec{x} \in \mathbb{R}^n} \vec{a} \cdot \vec{x} = +\infty.$$

That is, there is no maximum and by choosing $\vec{x}$, the value $\vec{a} \cdot \vec{x}$ can be as large as possible.

B) For given $\vec{c}, \vec{w} \in \mathbb{R}^n$, it must be true that

$$\max_{\vec{x} \in \mathbb{R}^n} [\vec{c} \cdot \vec{x} + \vec{w} \cdot \vec{x}] \geq \max_{\vec{x} \in \mathbb{R}^n,\ \vec{x} \geq \vec{0}} [\vec{c} \cdot \vec{x}].$$

C) A and B are both wrong

$$\max_{\vec{x}\in\mathbb{R}^n} \left[\vec{a}^T\vec{x}\right] = \begin{cases} 0 & \text{if } \vec{a}=\vec{0} \\ +\infty & \text{if } \vec{a}\neq\vec{0} \end{cases}$$

$$\left(\begin{array}{l} \text{Can choose } \vec{x} = \vec{a}\,t \quad t>0. \\ \vec{a}^T\vec{x} = \underbrace{|\vec{a}|^2}_{\neq 0.} t \longrightarrow \infty \quad \text{as } t\to\infty \\ \text{if } \vec{a}\neq\vec{0}. \end{array} \right)$$

$$\min_{\vec{x}\in\mathbb{R}^n} \left[\vec{a}^T\vec{x}\right] = \begin{cases} 0 & \text{if } \vec{a}=\vec{0} \\ -\infty & \text{if } \vec{a}\neq\vec{0}. \end{cases}$$

$$\max_{\vec{x}\in\mathbb{R}^n} \left(\vec{c}^T\vec{x} + \vec{w}^T\vec{x}\right) = \max_{\vec{x}\in\mathbb{R}^n} \left(\vec{c}+\vec{w}\right)^T\vec{x}$$

$$= \begin{cases} 0 & \text{if } \vec{c}+\vec{w}=\vec{0} \\ +\infty & \text{if } \vec{c}+\vec{w}\neq\vec{0}. \end{cases}$$

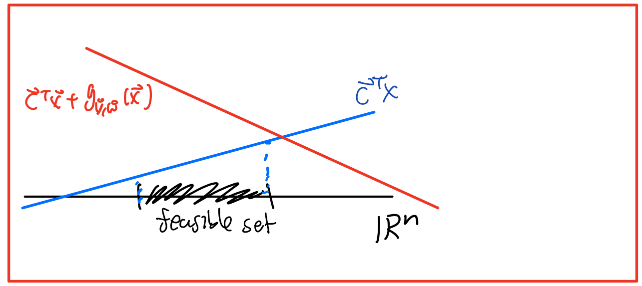
e.g. $n=1$, if $c=1$, $w=-1$.

then $\max\limits_{x\in\mathbb{R}} (c+w)\,x = 0$ while $\max\limits_{x\geq 0} c\,x \underset{\text{when } c=1}{=} \max\limits_{x\geq 0} x = +\infty.$

Back to our main discussion:

See $\vec{c}^T\vec{x} + g_{\vec{y},\vec{w}}(\vec{x}) = c^T\vec{x} + \vec{y}^T(\vec{b} - A\vec{x}) + w^T\vec{x}$

$$= \underbrace{\vec{y}^T\vec{b}}_{\vec{b}^T\vec{y}} + \underbrace{(\vec{c}^T - \vec{y}^TA + \vec{w}^T)\vec{x}}_{(\vec{c}+\vec{w}-A^T\vec{y})^T\vec{x}}$$

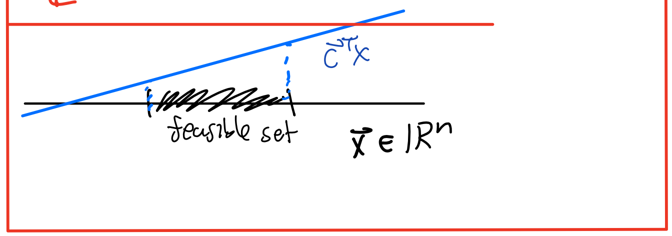


note $(\vec{a}+\vec{b})^T = \vec{a}^T + \vec{b}^T$

$(\vec{y}^TA)^T = A^T(\vec{y}^T)^T = A^T\vec{y}$

$\vec{c}^T\vec{x} + g_{\vec{y},\vec{w}}(\vec{x}) = \vec{y}^T\vec{b}$ if $\vec{c}+\vec{w}-A^T\vec{y} = \vec{0}$

constant as a function of $\vec{x}$



So,

$$F(\vec{y},\vec{w}) = \max_{\vec{x}}\left[\vec{c}^T\vec{x} + g_{\vec{y},\vec{w}}(\vec{x})\right] = \vec{b}^T\vec{y} + \max_{\vec{x}}\left[(\vec{c}+\vec{w}-A^T\vec{y})^T\vec{x}\right]$$

Note that

$$\max_{\vec{x}}\left[\left[\vec{c}+\vec{w}-A^T\vec{y}\right]^T\vec{x}\right] = \begin{cases} 0 & \text{if } \vec{c}+\vec{w}-A^T\vec{y} = \vec{0} \\ +\infty & \text{otherwise.} \end{cases}$$

can choose any $\vec{x}$

So,

$$\boxed{F(\vec{y},\vec{w}) = \begin{cases} \vec{b}^Ty & \text{if } \vec{c}+\vec{w}-A^T\vec{y} = \vec{0} \\ +\infty & \text{otherwise} \end{cases}}$$

When we minimize $\bar{F}(\vec{y},\vec{w})$,

Therefore.

$$\boxed{\begin{array}{ll} \text{minimize} & F(\vec{y},\vec{w}) \\ \text{subject to} & \vec{y},\vec{w} \geq \vec{0}. \end{array}} \iff \boxed{\begin{array}{ll} \text{minimize} & \vec{b}^T\vec{y} \\ \text{subject to} & \vec{c}+\vec{w}-A^T\vec{y}=\vec{0} \\ & \vec{y},\vec{w} \geq \vec{0}. \end{array}}$$

* In all those they have the same minimum objective value.

$\left\{ \begin{array}{l} \vec{w} = A^T\vec{y}-\vec{c} \\ \vec{w} \geq \vec{0} \end{array} \right\}$ is equivalent to $A^T\vec{y}-\vec{c} \geq 0$

note $\vec{w}$ is the slack variable to $A^T\vec{y} \geq \vec{c}$.

$$\iff \boxed{\begin{array}{ll} \text{minimize} & \vec{b}^T\vec{y} \\ \text{subject to} & A^T\vec{y} \geq \vec{c} \\ & \vec{y} \geq \vec{0}. \end{array}}$$

Note.
This is the dual problem of

$$\begin{array}{ll} \text{maximize} & \vec{c}^T\vec{x} \\ \text{subject to} & A\vec{x} \leq \vec{b},\ \vec{x} \geq \vec{0} \end{array}$$

- All these three are equivalent.

(giving the same optimal solution
and the same optimal objective value).

Lec 10. TuTh.