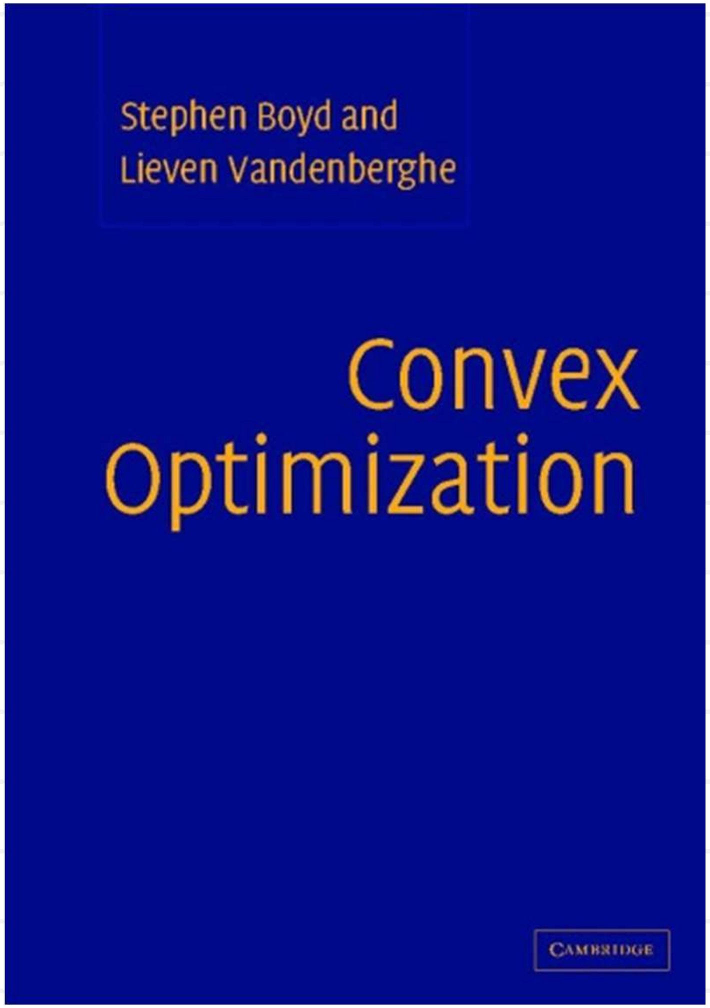
Lecture 14: Lagrange duality

Reading



Section 5

Lagrange duality in optimization

$$\min \; f_0(x)$$
$$\text{s.t.:} \; f_i(x) \leq 0$$
$$h_i(x) = 0$$

Primal problem

Dual problem

$$\max \; g(\lambda, \nu)$$
$$\text{s.t.:} \; k_i(\lambda, \nu) \leq 0$$
$$l_i(\lambda, \nu) = 0$$

The dual problem is:

- Convex (concave)
- Sometimes simpler
- Always a lower-bound
- Sometimes a tight one
- Solution to the dual may lead to a good solution of the primal
- A large number of optimization methods are based on Lagrange dual

The Lagrangian

$$\min f_0(x)$$
$$\text{s.t. } f_i(x) \le 0, \quad i=1..m$$
$$h_i(x) = 0, \quad i=1...p$$
$x \in D$

$p^* = \min_{x \in D} f_0(x)$

The *Lagrangian*:

$$L(x,\lambda,\nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$$

Observation: suppose $x \in D$ and $\lambda \geq 0$. Then

$$L(x,\lambda,\nu) \le f_0(x)$$

Consequence: $\forall \lambda \geq 0, \nu$

$$\min_x L(x,\lambda,\nu) \le \min_{x \in D} L(x,\lambda,\nu) \le \min_{x \in D} f_0(x)$$

The Lagrange dual function

Consequence: $\forall \lambda \geq 0, \nu$

$$\min_x L(x,\lambda,\nu) \leq \min_{x\in D} L(x,\lambda,\nu) \leq \min_{x\in D} f_0(x) = p^*$$

The dual function:

$$g(\lambda,\nu) = \underset{x}{\overset{(\inf)}{\min}}\, L(x,\lambda,\nu)$$

Theorem (*weak duality*):

$$\max_{\lambda\geq 0,\nu} g(\lambda,\nu) \leq \min_{\substack{f_i(x)\leq 0\\ h_i(x)=0}} f_0(x)$$

Primal and dual programs

Primal program *Dual program*

$$p^* = \min f_0(x) \geq d^* = \max g(\lambda, \nu)$$

$$\text{s.t.: } f_i(x) \leq 0 \qquad \text{s.t.: } \lambda \geq 0$$

$$h_i(x) = 0$$

$$g(\lambda, \nu) = \min_x L(x, \lambda, \nu)$$

- Dual has a "simpler" domain
- Dual objective is hard to compute
- In many important cases, however, it can be simplified analytically

LP duality

$$\begin{aligned} &\min\ c^T x \\ &\text{s.t. } Ax = b \\ &\quad\ \ x \geq 0 \end{aligned}$$

Lagrangian:

$$L(x,\lambda,\nu) = c^T x - \lambda^T x + \nu^T (b - Ax)$$

$$g(\lambda,\nu) = \min_x L(x,\lambda,\nu) =$$

$$= \min_x \left(c^T - \lambda^T - \nu^T A\right) x + \nu^T b$$

Dual program:

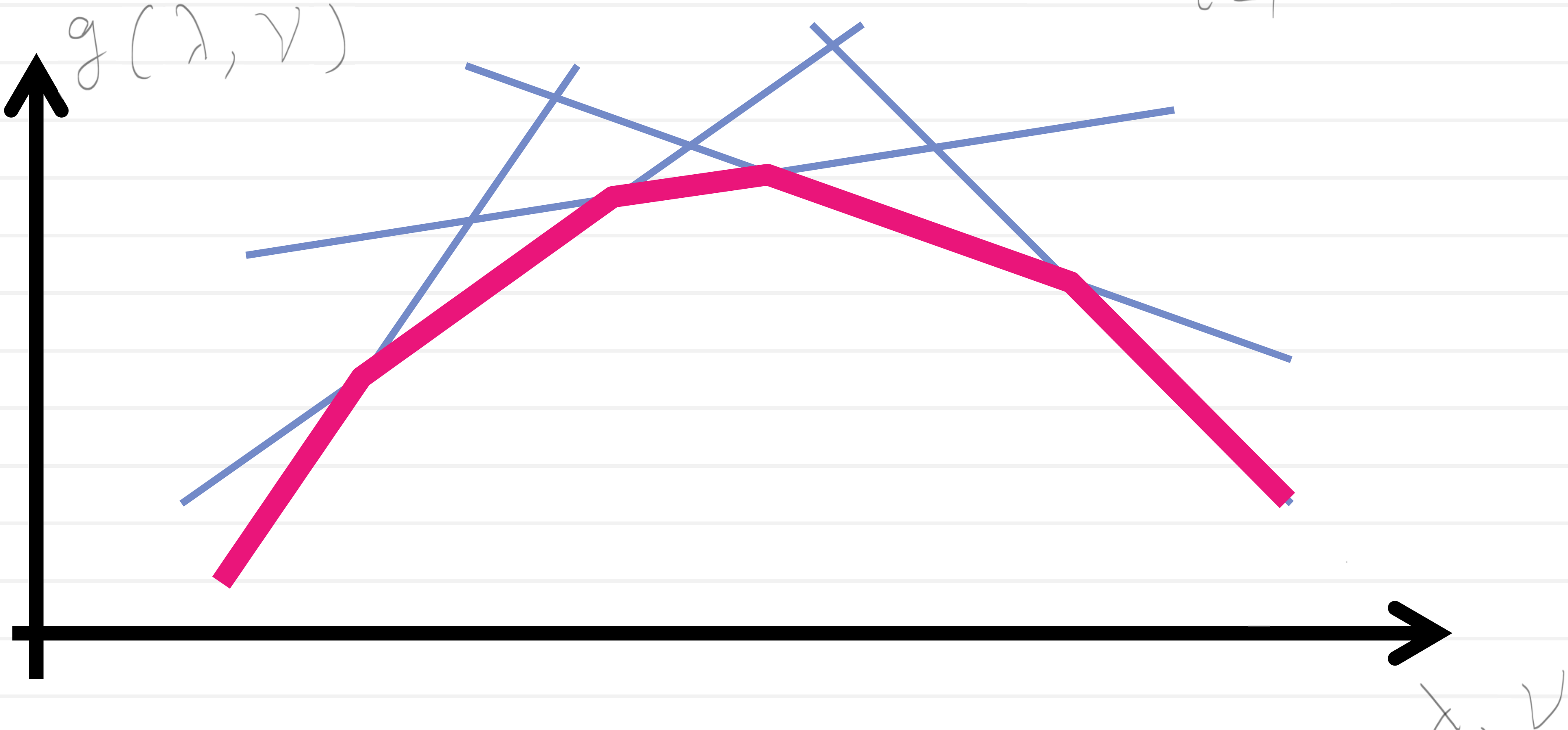
$$\begin{aligned} &\max\ \nu^T b \\ &\text{s.t. } \nu^T A + \lambda^T = c^T \\ &\qquad\quad \lambda \geq 0 \end{aligned}$$

$$\begin{aligned} &\max_\nu\ b^T \nu \\ &\text{s.t.: } A^T \nu \leq c \end{aligned}$$

$$\begin{aligned} &\min\ -b^T y \\ &\text{s.t.: } A^T y \leq c \end{aligned}$$

Dual is concave

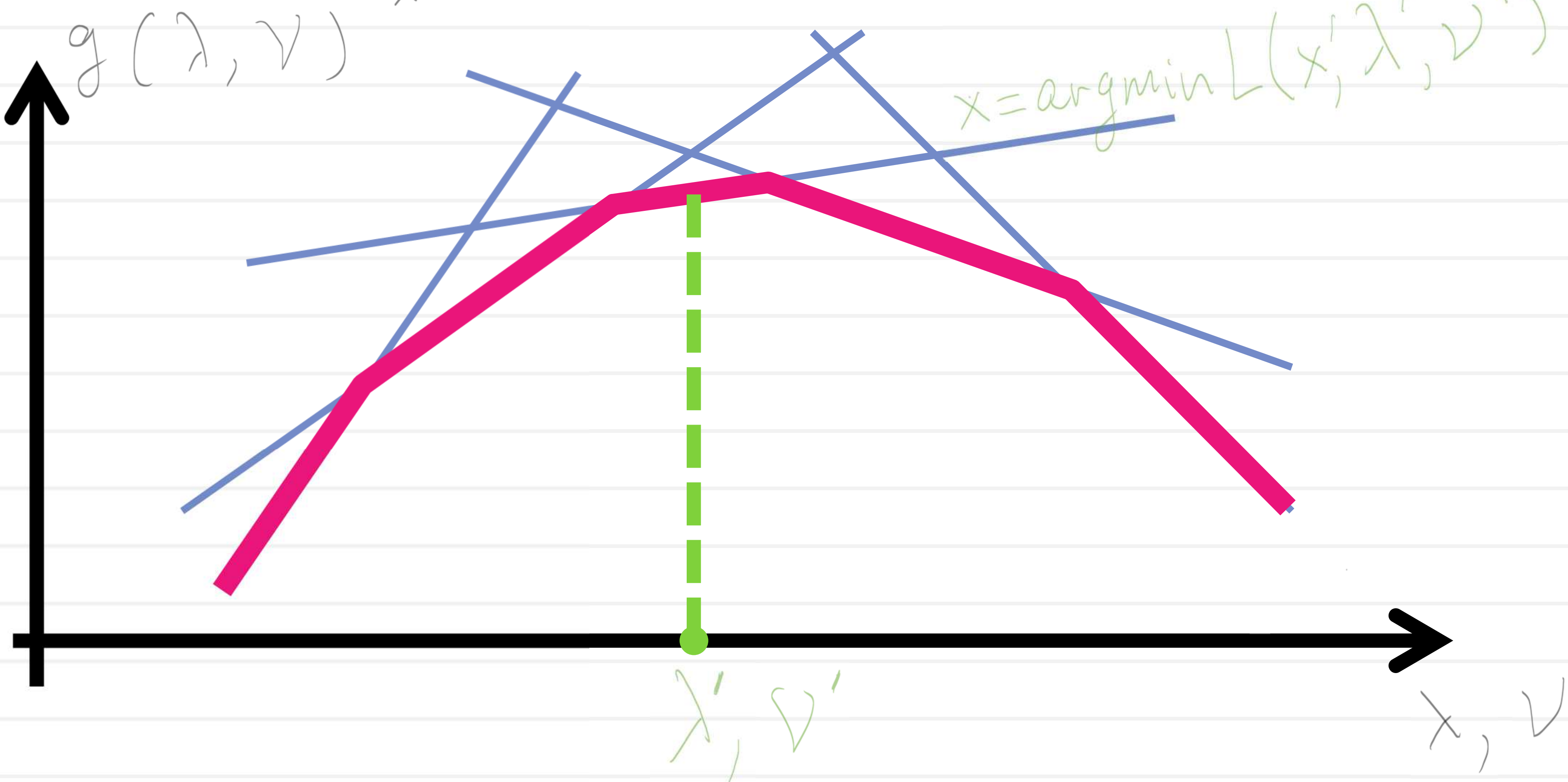
$$g(\lambda, \nu) = \min_x \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right)$$



- Pointwise minimum of a family of linear functions
- Dual program always leads to convex optimization

Dual is concave

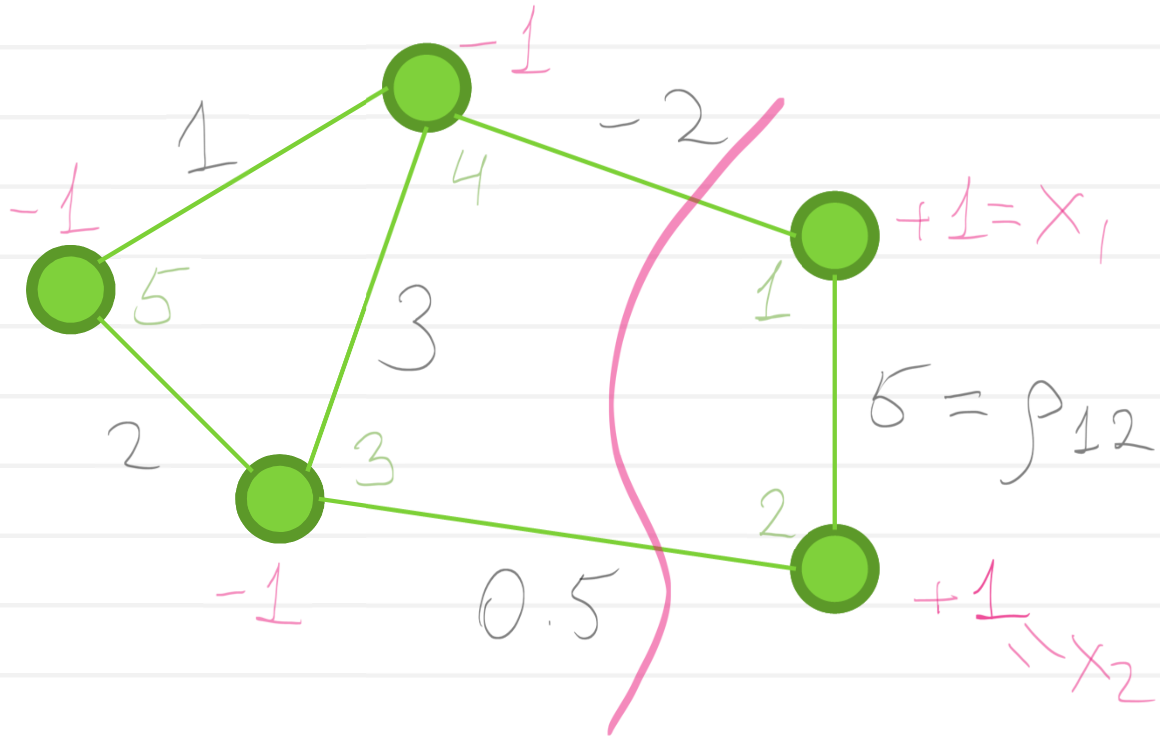
$$g(\lambda,\nu) = \min_x L(x,\lambda,\nu)$$



$$g(\lambda,\nu) \le g(\lambda',\nu') + f(x')^T(\lambda-\lambda') + h(x')^T(\nu-\nu')$$

$\nabla g = (f_i(x), h_i(x))$ *- (sub)gradient of g*

Graph partition



Task: split/partition the graph into two parts to minimize the partition weight

We can express the partition cost as: $\sum \rho_{ij} (x_i - x_j)^2 \cdot \frac{1}{4}$

We get the following integer program:

$$x^T W x \to \min$$

$$\text{s.t.:} \quad x_i \in \{-1; 1\}$$

Graph partition: the dual

$$\min \; x^T W x$$
$$\text{s.t.} \quad x_i^2 = 1 \quad (\text{i.e. } x_i = \pm 1)$$

$$L(x, \nu) = x^T W x + \nu^T (x_i^2 - 1)$$

$$g(\nu) = \min_x \left(x^T (W + \operatorname{diag} \nu) x - \sum_i \nu_i \right)$$

$$\boxed{\begin{aligned} &\min_{\nu} \; \sum_i \nu_i \\ &\text{s.t.}: \quad W + \operatorname{diag}(\nu) \succcurlyeq 0 \end{aligned}}$$

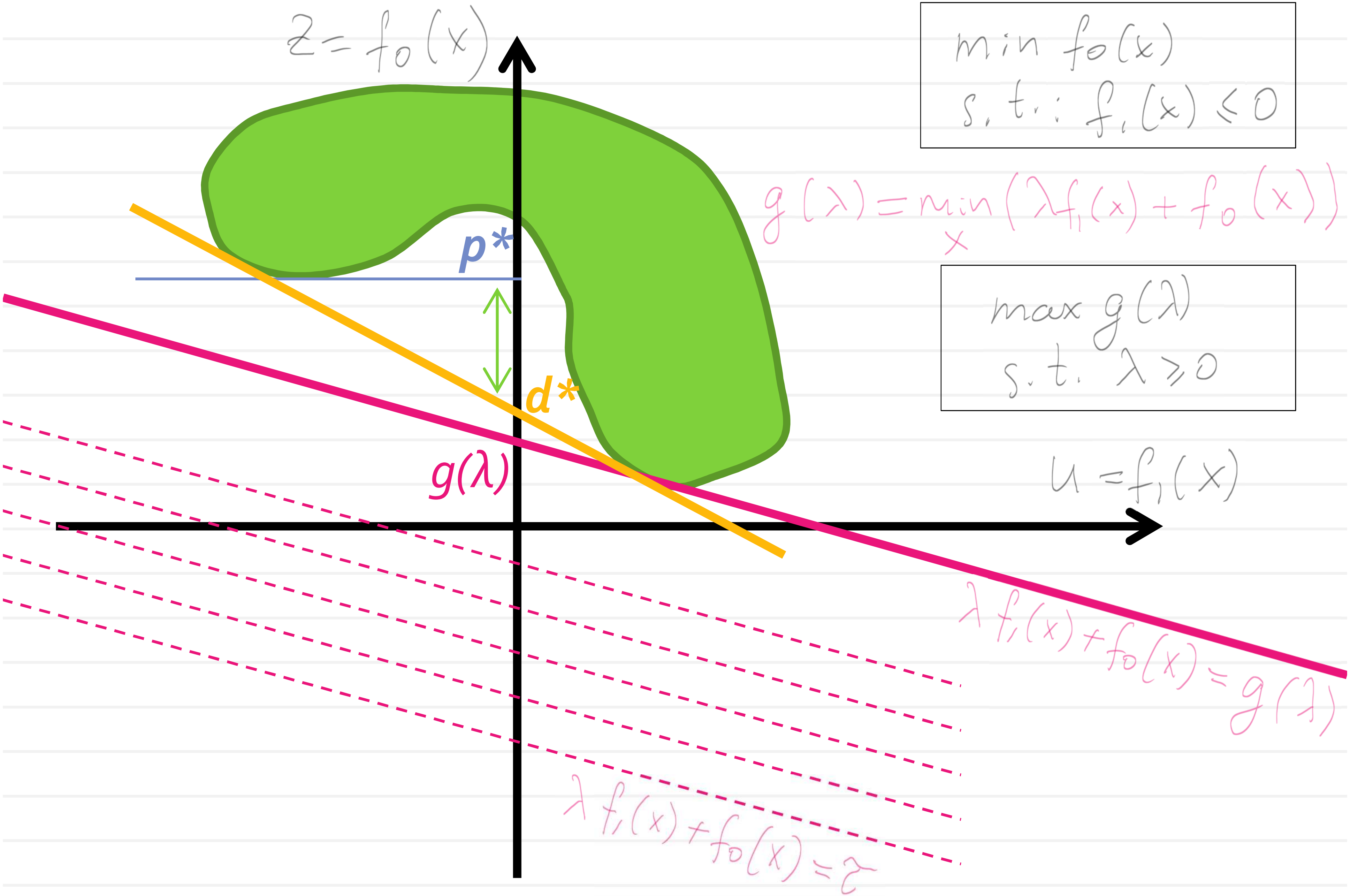
Semidefinite program (convex!)

Geometric intuition: weak duality

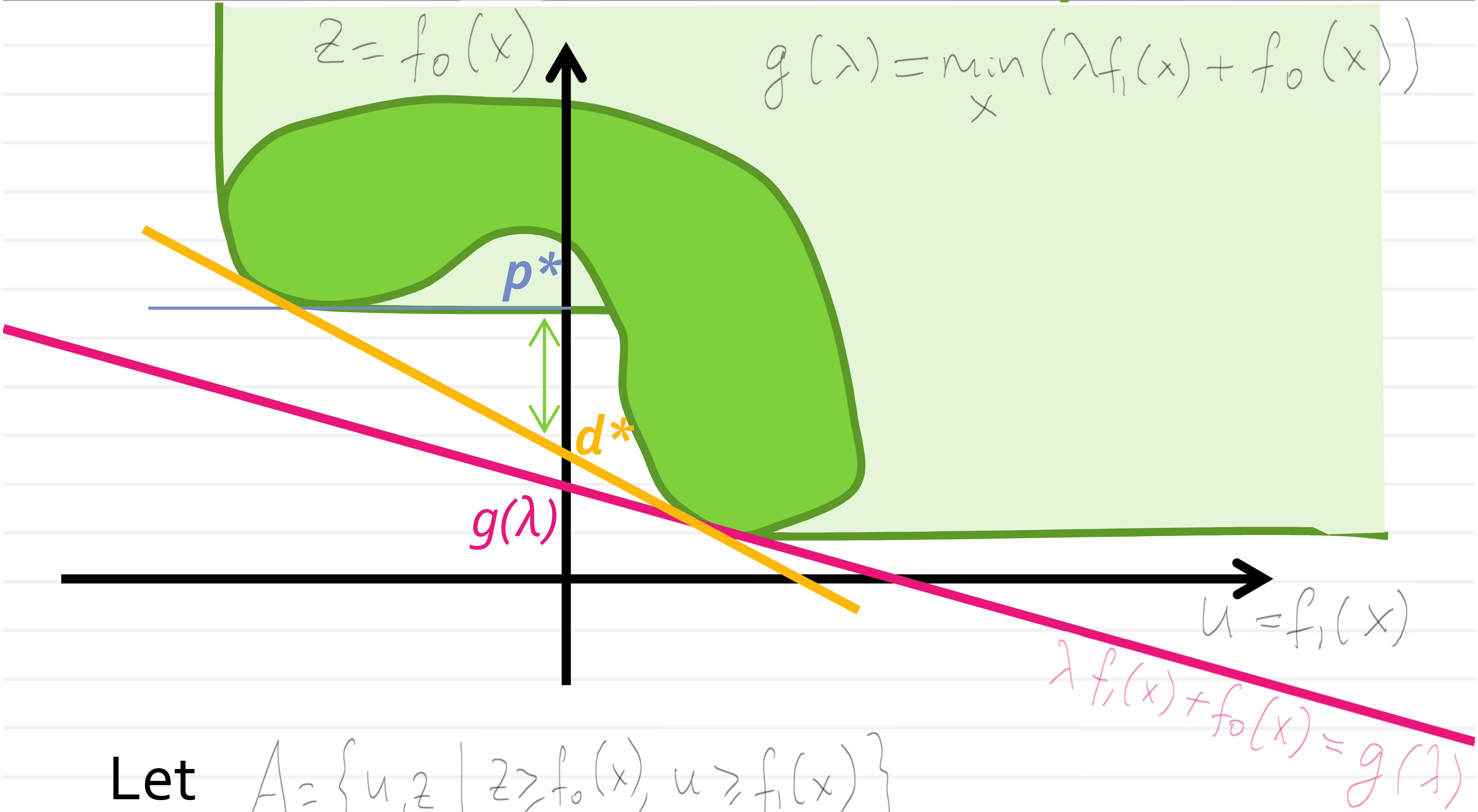
$$\min f_0(x) \quad \text{s.t.: } f_1(x) \leq 0$$

$$g(\lambda) = \min_x \left(\lambda f_1(x) + f_0(x)\right)$$

$$\max g(\lambda) \quad \text{s.t. } \lambda \geq 0$$



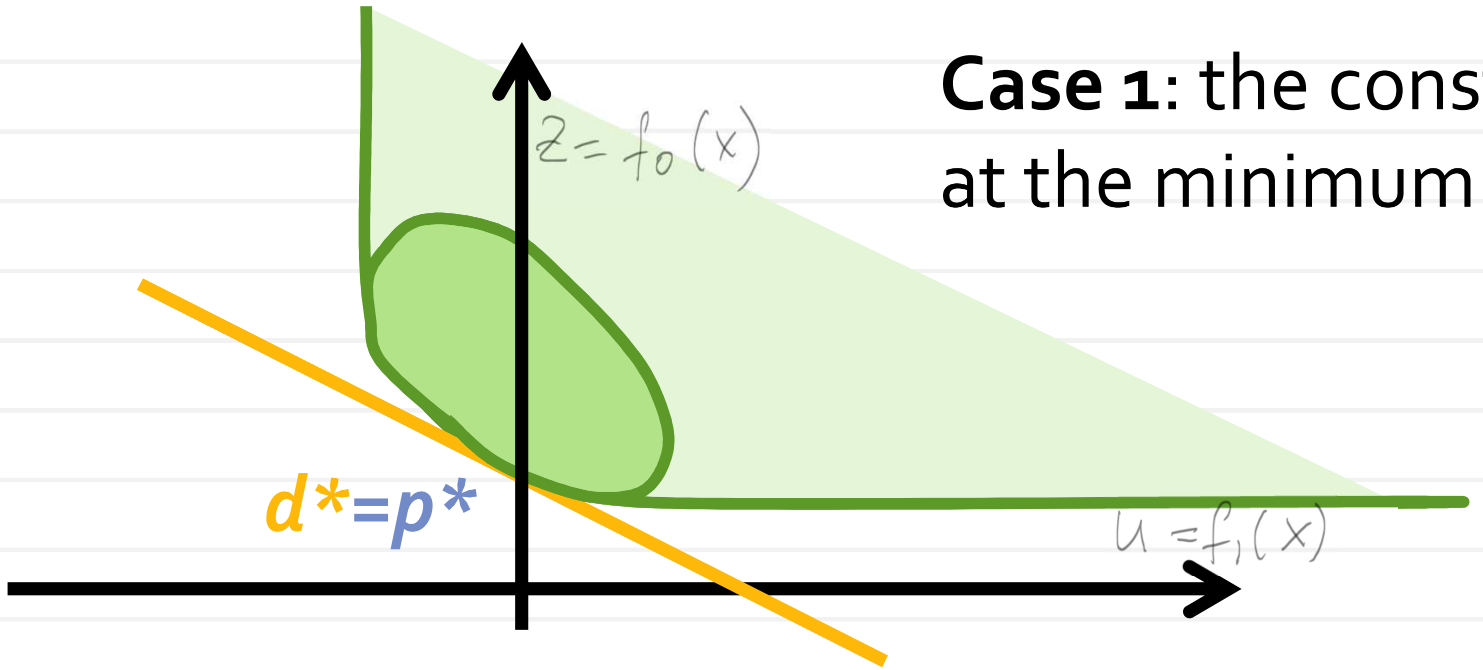
Geometric intuition: weak duality



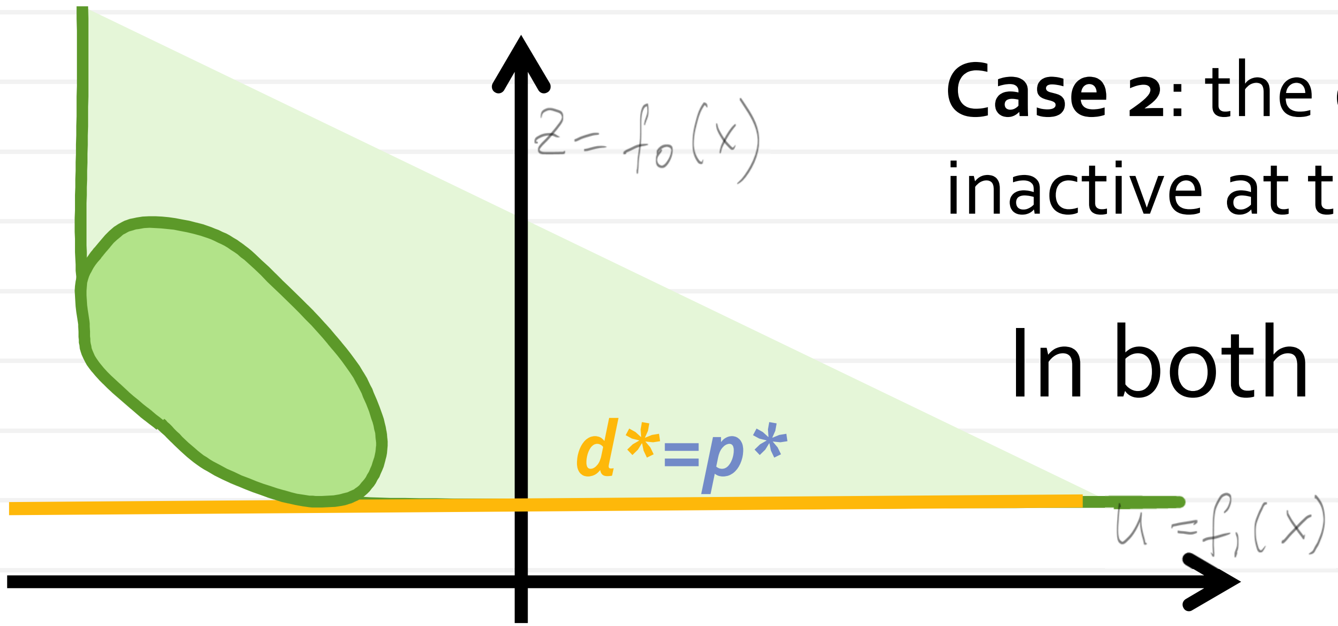
Let $A = \{u, z \mid z \geq f_0(x), u \geq f_1(x)\}$

This new set (lightgreen) is convex for convex problems!

Geometric intuition: convex case



Case 1: the constraint is active at the minimum



Case 2: the constraint is inactive at the minimum

In both cases, d*=p*!

Strong duality

$$\min f_0(x)$$
$$\left.\begin{array}{l} \text{s.t.: } f_i(x) \le 0 \\ \quad\quad Ax = b \end{array}\right\} D$$

If the primal problem is convex and its domain has non-empty relative interior

$$\exists x, N(x): N(x) \cap \{Ax=b\} \in D$$

then the optimal primal value equals the optimal dual value (duality gap is zero).

LP duality

$$\min \; c^T x$$
$$\text{s.t. } Ax = b$$
$$x \geq 0$$

$$||$$

$$\max_{\nu} \; b^T \nu$$
$$\text{s.t.: } A^T \nu \leq c$$

For linear programs the strong duality always holds, as long as at least one of the problems is feasible

Economic interpretation

$\min f_0(x)$
s.t. $f_i(x) \le 0$

x is how we operate,
$f_o(x)$ is our incurred cost,
$f_i(x)$ are constraints on how we can operate

Let us assume that we pay a penalty/get reward λ_i for any unit violation/underuse of the ith constraint.

Our new cost: $L(x,\lambda) = f_0(x) + \sum_i \lambda_i f_i(x)$

If we operate optimally, the cost will be: $g(\lambda)$

Strong duality: there exist a set of *shadow* prices λ^* so that violating/underusing constraints do not give us any extra profit. $d^* = g(\lambda^*) = f_0(x^*) = p^*$

Recap: Lagrange duality in optimization

$$\min \; f_0(x)$$
$$\text{s.t.:} \quad f_i(x) \le 0$$
$$h_i(x) = 0$$

Primal problem

Dual problem

$$\max \; g(\lambda, \nu)$$
$$\text{s.t.:} \quad k_i(\lambda, \nu) \le 0$$
$$l_i(\lambda, \nu) = 0$$

The dual problem is:

- Convex (concave)
- Sometimes simpler
- Always a lower-bound
- Sometimes a tight one
- Solution to the dual may lead to a good solution of the primal
- A large number of optimization methods are based on Lagrange dual