The Lagrange dual function

- Lagrangian $L: \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$

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

- Lagrange dual function $g: \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$

$$g(\lambda, \nu) = \inf_{x \in D} L(x, \lambda, \nu)$$

$$= \inf_{x \in D} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right)$$

pointwise minimum of affine function of (λ, ν) → concave

The Lagrange dual function and conjugate functions

The conjugate function and Lagrange dual function are closely related.

Ex 1

$$\boxed{\begin{array}{ll} \min. & f(x) \\ \text{s.t.} & x = 0 \end{array}}$$

$$L(x, \nu) = f(x) + \nu^T x$$

$$g(\nu) = \inf_x \left(f(x) + \nu^T x \right)$$

$$= -\sup_x \left(-\nu^T x - f(x) \right)$$

$$= -f^*(-\nu)$$

Ex 2

$$\boxed{\begin{array}{ll} \min. & f_0(x) \\ \text{s.t.} & Ax \preceq b \\ & Cx = d \end{array}}$$

$$g(\lambda, \nu) = \inf_x \left(f_0(x) + \lambda^T (Ax - b) + \nu^T (Cx - d) \right)$$

$$= -b^T \lambda - d^T \nu + \inf_x \left(f_0(x) + (A^T \lambda + C^T \nu)^T x \right)$$

$$= -b^T\lambda - d^T\nu - \sup_x \left((A^T\lambda + C^T\nu)^T x - f_0(x) \right)$$

$$= -b^T\lambda - d^T\nu - f^*(-A^T\lambda - C^T\nu)$$

(iii) The Lagrange dual problem

$$\boxed{\begin{array}{ll} \max. & g(\lambda, \nu) \\ \text{s.t.} & \lambda \succeq 0 \end{array}}$$

Convex optimization problem!

(iv) Duality

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

$d^* = p^* \rightarrow$ strong duality

(d^* ... dual optimal value

p^* ... primal optimal value)

(IV) Complementary slackness

Suppose that the strong duality holds, then

$$
\begin{aligned}
f_0(x^*) &= g(\lambda^*, \nu^*) \\
&= \inf_x \left(f_0(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{i=1}^{p} \nu_i^* h_i(x) \right) \\
&\leq f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* f_i(x^*) + \nu_i^* h_i(x^*) \\
&\leq f_0(x^*)
\end{aligned}
$$

$$\Longrightarrow \quad \lambda_i^* f_i(x^*) = 0$$

Complementary slackness

(V) KKT Conditions

Assume $f_0 \cdots f_m$, $h_1 \cdots h_p$ are differentiable and strong duality holds, then besides the constraints we have

$$\nabla_x L(x, \lambda^*, \nu^*)\Big|_{x=x^*} = 0$$

So

$$f_i(x^*) \le 0$$

$$h_i(x^*) = 0$$

$$\lambda_i^* \ge 0$$

natural

$$\lambda_i^* f_i(x^*) = 0$$

← complementary slackness

$$\nabla f_0(x) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x) + \sum_{i=1}^{p} \nu_i^* \nabla h_i(x)\Big|_{x=x^*} = 0$$

↑ differentiation

KKT conditions

Ⅲ Examples of dual problem

① Least-square solution of linear equations

$$\boxed{\begin{array}{ll} \min. & x^T x \\ \text{s.t.} & Ax = b \end{array}}$$

$$g(\nu) = \inf_x \left(\underbrace{x^T x + \nu^T(Ax-b)}_{\text{convex quadratic func of } x} \right)$$

$$\nabla_x \left(x^T x + \nu^T(Ax-b)\right) = 2x + A^T\nu = 0$$

$$\Rightarrow \quad x = -\frac{1}{2}A^T\nu$$

$$= -\frac{1}{4}\nu^T A A^T \nu - \nu^T b$$

Associated dual problem is

$$\boxed{\text{max.} \quad -\frac{1}{4}\nu^T A A^T \nu - \nu^T b}$$

② LP

$$\boxed{\begin{array}{ll} \text{min.} & c^T x \\ \text{s.t.} & Ax = b \\ & x \succeq 0 \end{array}}$$

$$g(\lambda,\nu) = \inf_x \left(c^T x + \nu^T(Ax-b) - \lambda^T x \right)$$

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

$$= \begin{cases} -b^T\nu & \text{if} \quad c - \lambda + A^T\nu = 0 \\ -\infty & \text{otherwise} \end{cases}$$

$\leftarrow$ dual infeasible

Associated dual problem is

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

$\downarrow$

$$\boxed{\begin{aligned} \text{max.} \quad & -b^T\nu \\ \text{s.t.} \quad & c + A^T\nu \succeq 0 \end{aligned}}$$

③ QCQP

$$\boxed{\begin{aligned} \text{min.} \quad & \frac{1}{2} x^T P_0 x + q_0^T x + r_0 \\ \text{s.t.} \quad & \frac{1}{2} x^T P_i x + q_i^T x + r_i \leq 0 \end{aligned}}$$

with $P_0 \in S^n_{++}$, $P_i \in S^n_+$

$$g(\lambda) = \inf_x \left(\frac{1}{2} x^T P_0 x + q_0^T x + r_0 + \sum_{i=1}^m \lambda_i \left(\frac{1}{2} x^T P_i x + q_i^T x + r_i \right) \right)$$

$$= \inf_x \left(\frac{1}{2} x^T \left(P_0 + \sum_{i=1}^m \lambda_i P_i \right) x + \left(q_0 + \sum_{i=1}^m \lambda_i q_i \right)^T x + \left(r_0 + \sum_{i=1}^m \lambda_i r_i \right) \right)$$

$$= \inf_x \left(\frac{1}{2} x^T P(\lambda) x + q(\lambda)^T x + r(\lambda) \right)$$

$$\left(\begin{aligned} P(\lambda) &= P_0 + \sum_{i=1}^m \lambda_i P_i \\ q(\lambda) &= q_0 + \sum_{i=1}^m \lambda_i q_i \\ r(\lambda) &= r_0 + \sum_{i=1}^m \lambda_i r_i \end{aligned} \right)$$

$$= -\frac{1}{2} q^T P^{-1} q + r$$

dual problem is

$$\boxed{\begin{array}{ll} \max. & -\frac{1}{2} q(\lambda)^T P^{-1}(\lambda) q(\lambda) + r(\lambda) \\ & \lambda \geq 0 \end{array}}$$

④ Entropy maximization

$$\boxed{\begin{array}{ll} \min. & \sum_{i=1}^{n} x_i \log x_i \\ \text{s.t.} & Ax \preceq b \\ & \mathbb{1}^T x = 1 \end{array}}$$

☆ Use the conjugate function!

Consider the following problem:

$$\boxed{\begin{array}{ll} \min. & f_0(x) \\ \text{s.t.} & Ax \leq b \\ & Cx = d \end{array}}$$

$$g(\lambda, \nu) = \inf_x \left(f_0(x) + \lambda^T(Ax-b) + \nu^T(Cx-d) \right)$$

$$= -b^T\lambda - d^T\nu - \sup_x \left(-f_0(x) + (-A^T\lambda - C^T\nu)^T x \right)$$

$$= -b^T\lambda - d^T\nu - f_0^*(-A^T\lambda - c^T\nu)$$

Now the conjugate of the function

$$f_0(x) = \sum_{i=1}^n x_i \log x_i$$

is

$$f_*(y) = \sum_{i=1}^n e^{y_i - 1}$$

therfore

$$g(\lambda, \nu) = -b^T\lambda - \nu - \sum_{i=1}^n e^{-a_i^T\lambda - \nu - 1}$$

$$= -b^T\lambda - \nu - e^{-\nu-1} \sum_{i=1}^n e^{-a_i^T\lambda}$$

Therefore, associated dual function is:

$$\boxed{\begin{array}{ll} \max. & -b^T\lambda - \nu - e^{-\nu-1} \sum_{i=1}^n e^{-a_i^T\lambda} \\ \text{s.t.} & \lambda \succeq 0 \end{array}}$$

- maximizing over ν

$$\nabla_\nu \left(-b^T\lambda - \nu - e^{-\nu-1} \sum_{i=1}^{n} e^{-a_i^T\lambda}\right)$$

$$= -1 + e^{-\nu-1} \sum_{i=1}^{n} e^{-a_i^T\lambda} = 0$$

$$\Rightarrow \nu = \log\left(\sum_{i=1}^{n} e^{-a_i^T\lambda}\right) - 1$$

$$\boxed{\begin{aligned} \max.\quad & -b^T\lambda - \log\left(\sum_{i=1}^{n} e^{-a_i^T\lambda}\right) \\ \text{s.t.}\quad & \lambda \succeq 0 \end{aligned}}$$

Solving the primal problem via the dual

If strong duality holds, then any primal optimal point is also a minimizer of $L(x, \lambda^*, \nu^*)$, i.e. the solution of

$$\boxed{\min.\quad f_0(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{i=1}^{p} \nu_i^* h_i(x)}$$

is the primal optimal.

The perturbed problem

$$\boxed{\begin{aligned} \min.\quad & f_0(x) \\ \text{s.t.}\quad & f_i(x) \le u_i \\ & h_i(x) = v_i \end{aligned}}$$

$$\left(\begin{aligned} u_i > 0 \quad &\rightarrow \quad \text{relaxed} \\ u_i < 0 \quad &\rightarrow \quad \text{tightened} \end{aligned} \right)$$

We define the optimal value of the perturbed problem as :

$$p^*(u, v) = \inf \left\{ f_0(x) \,\middle|\, \exists x \in D,\ f_i(x) \le u_i,\ h_i(x) = v_i \right\}$$

Then we get the following fact :

<u>When the original problem is convex, then function p^* is a convex function of u and v</u>

⑪ A global inequality

We assume that strong duality holds,
(Original problem is convex and Slater's condition is satisfied)

by strong duality,

$$\underbrace{p^*(0,0)}_{\text{primal optimal}} = \underbrace{g(\lambda^*, \nu^*)}_{\text{dual optimal}}$$

$$\leq f_0(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{i=1}^{p} \nu_i^* h_i(x)$$

$$\Big\downarrow \begin{array}{l} f_i(x) \leq u_i \\ h_i(x) = v_i \end{array}$$

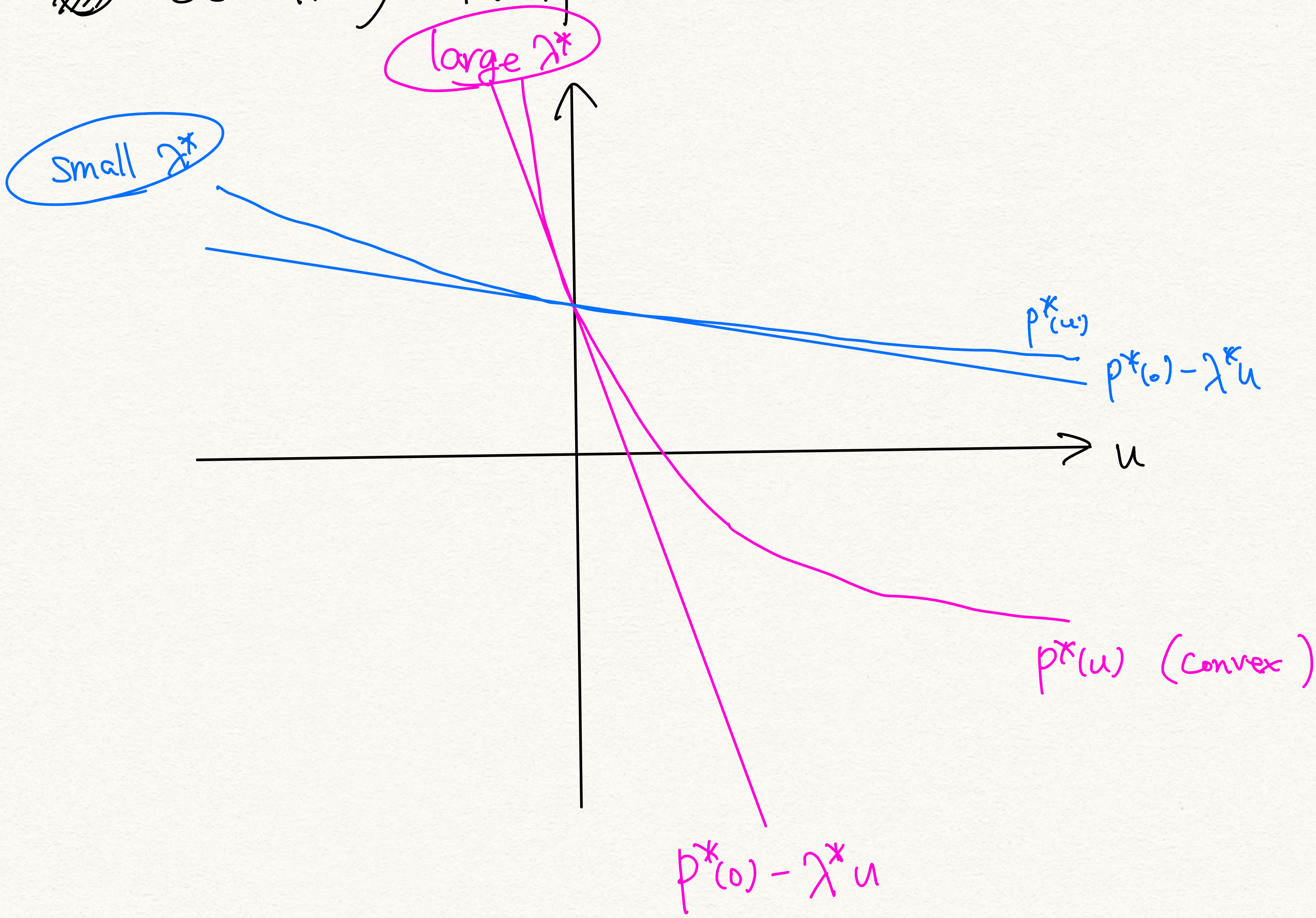
$$\leq f_0(x) + \lambda^{*T} u + \nu^{*T} v$$

(x : any feasible point for the perturbed problem, i.e. $x \in D$, $f_i(x) \leq u_i$, $h_i(x) = v_i$)

$$\Leftrightarrow \quad f_0(x) \geq p^*(0,0) - \lambda^{*T} u - \nu^{*T} v$$

$$\Rightarrow \quad \underbrace{p^*(u,v) \geq p^*(0,0) - \lambda^{*T} u - \nu^{*T} v}_{\text{global inequality}}$$

Sensitivity interpretation



Introduce constraints

Ex

$$\boxed{\text{min.} \quad f_0(Ax+b)}$$

← no constraints

$\downarrow$

$$\boxed{\begin{array}{ll} \text{min.} & f_0(y) \\ \text{s.t.} & y = Ax+b \end{array}}$$

$$g(\nu) = \inf_{x,y} \left(f_0(y) + \nu^T (Ax + b - y) \right)$$

$$= b^T\nu + \inf_{x,y} \left((A^T\nu)^T x - \nu^T y + f_0(y) \right)$$

if $A^T\nu \neq 0$, then $g(\nu) = -\infty$

so $A^T\nu = 0$

$$= b^T\nu + \inf_{y} \left(-\nu^T y + f_0(y) \right)$$

$$= b^T\nu - \sup_{y} \left(\nu^T y - f_0(y) \right)$$

$$= b^T\nu - f_0^*(\nu)$$

Therefore, the dual problem is :

$$\boxed{\begin{aligned} \max.\quad & b^T\nu - f_0^*(\nu) \\ \text{s.t.}\quad & A^T\nu = 0 \end{aligned}}$$