Convex Optimization

Ex1 $\min_x c^T x$ s.t $Ax-b=0$ (P), $-x\leq 0$

$\min_y -b^T y$ s.t $A^T y - c \leq 0$ (D)

} rewritten in standard form

1) $L^P(x,\lambda,\mu) = c^T x - \lambda^T x + \mu^T(Ax-b)$ is convex in x

The dual function is

$$g^P(\lambda,\mu) = \inf_x \{c^T x - \lambda^T x + \mu^T(Ax-b)\}$$

$$\frac{\partial L^P}{\partial x}(x,\lambda,\mu) = c - \lambda + A^T\mu = 0 \quad \text{(constant slope)}$$

$$L^P(x,\lambda,\mu) = (\lambda - A^T\mu)^T x - \lambda^T x + \mu^T A x - \mu^T b$$
$$= (-\mu^T A + \lambda^T)x - \lambda^T x + \mu^T A x - \mu^T b$$
$$= \cancel{-\mu^T A x} + \cancel{\lambda^T x} - \cancel{\lambda^T x} + \cancel{\mu^T A x} - \mu^T b = -\mu^T b$$

$$\Rightarrow g^P(\lambda,\mu) = \begin{cases} -\mu^T b & \text{if } c-\lambda+A^T\mu = 0 \\ -\infty & \text{o.w} \end{cases}$$

Then the dual problem is $\max -\mu^T b$ st $\lambda \geq 0$

Which is the problem $\boxed{\max_\mu -\mu^T b \quad c + A^T\mu \geq 0}$ since $\lambda = A^T\mu + c$

2) $L^D(y,\lambda) = -b^T y + \lambda^T(A^T y - c)$

The dual function is $g^P(\lambda) = \inf_y \{-b^T y + \lambda^T(A^T y - c)\}$

$$\frac{\partial L^D}{\partial y}(y,\lambda) = -b + A\lambda = 0 \quad \rightarrow b = A\lambda$$

$$g^D(\lambda) = \cancel{-(A\lambda)^T y} + \cancel{\lambda^T A^T y} - \lambda^T c = -\lambda^T c$$

Then $g^D(\lambda) = \begin{cases} -\lambda^T c & \text{if } b = A\lambda \\ -\infty & \text{o.w} \end{cases}$

And the dual problem is $\boxed{\begin{array}{l} \max_\lambda -\lambda^T c \\ \text{st } \lambda \geq 0 \\ b - A\lambda = 0 \end{array}}$

Escaneado con CamScanner

Convex Optimization

Ex 1

3) $\min_{x,y} c^Tx - b^Ty$ st $\left.\begin{array}{l} Ax-b=0 \\ -x \leq 0 \\ A^Ty - c \leq 0 \end{array}\right\}$ standard form

This can be solved independently for x and y, and use part: 1) and 2)

(P) $\min_x c^Tx$ st $Ax-b=0$
$-x \leq 0$

(D) $\min_y -b^Ty$ st $A^Ty - c \leq 0$

with duals

(P^D) $\max_{\mu,\lambda_1} -\mu^T b$ st $\lambda_1 \geq 0$
$c - \lambda_1 + A^T\mu = 0$

(D^D) $\max_{\lambda_2} -\lambda_2^T c$ st $\lambda_2 \geq 0$
$b - A\lambda_2 = 0$

They can be joined in

$\max_{\mu,\lambda_1,\lambda_2} -\mu^T b - \lambda_2^T c$ st $\lambda_1 \geq 0$
$\lambda_2 \geq 0$
$c - \lambda_1 + A^T\mu = 0$
$b - A\lambda_2 = 0$

Since $\lambda_1 = c + A^T\mu$ we can write

$c + A^T\mu \geq 0$

$\min_{\mu,\lambda_2} \mu^T b + \lambda_2^T c$ s.t $c + A^T\mu \geq 0$
$b - A\lambda_2 = 0$
$\lambda_2 \geq 0$

Now we call $y = -\mu$ and $x = \lambda_2$ and we get the equivalent problem

$\min_{x,y} c^Tx - b^Ty$ st $A^Ty \leq c$

$Ax = b$

$x \geq 0$

This proves that the problem is self-dual

Convex Optimization

Ex1 4) Strong duality $\Rightarrow d^* = p^* = (x^*, y^*)$

As we saw before $\min_{x,y} \; c^T x - b^T y$ st $Ax = b$
$x \geq 0$
$A^T y \leq c$

can be solve by minimizing first over one variable and then over the other one (since constraints are indep)

Moreover, by solving (P) we get x^* and by solving (D) we get y^*, then the min is $c^T x^* - b^T y^*$.

b) Since strong duality holds $c^T x^* = -\mu^{*T} b$

Note that the dual of (P) is the same as (D) with a change of variable $y = -\mu$, thus $c^T x^* = b^T y^*$

Finally, the solution to the self-dual is $c^T x^* - b^T y^* = 0$

Convex Optimization

Ex 2| $\min_x \|Ax-b\|_2^2 + \|x\|_1$

1) $f(x) = \|x\|_1$

$f^*(y) = \sup_x (\underbrace{y^T x - \|x_1\|}_{g(x,y)})$

$g(x,y) = \sum_{i=1}^{d} y_i x_i - \sum_{i=1}^{d} |x_i| = \sum_{i=1}^{d} y_i x_i - |x_i|$

• If $y_i > 1$ for some i:

Let $x_i = \alpha$ and $x_j = 0 \ \forall j \neq i$, $\alpha > 0$

$g(x,y) = \alpha \underbrace{(y_i - 1)}_{>1} \xrightarrow{\alpha \to +\infty} +\infty \Rightarrow$ sup is $+\infty$

• If $y_i \leq 1 \ \forall i$:

• If $y_i < -1$ for some i:

Let $x_i = \alpha < 0$ and $x_j = 0 \ \forall j \neq i$

$g(x,y) = y_i x_i + x_i = x_i \underbrace{(y_i + 1)}_{<0} \xrightarrow{\alpha \to -\infty} +\infty \Rightarrow$ sup is $+\infty$

• If $y_i \geq -1 \ \forall i$:

Let $x = \alpha \mathbb{1}$ $\alpha > 0$ then we have

$g(x,y) = \alpha \left(\sum_{i=1}^{d} y_i - d \right)$

• If $\sum_{i=1}^{d} y_i - d > 0$:

$g(x,y) = \underbrace{\alpha}_{>0} \underbrace{\left(\sum_{i=1}^{d} y_i - d \right)}_{>0} \xrightarrow{\alpha \to +\infty} +\infty$

Let $x = \alpha \mathbb{1}$ $\alpha < 0$:

$g(x,y) = \alpha \left(\sum_{i=1}^{d} y_i + d \right)$

• If $\sum_{i=1}^{d} y_i + d < 0$: $g(x,y) \xrightarrow{\alpha \to -\infty} +\infty$

However note that the cases $\sum_{i=1}^{d} y_i > d$ and $\sum_{i=1}^{d} y_i < -d$ can't happen because $\underbrace{|y_i| \leq 1 \ \forall i}_{\text{This is } \|y\|_\infty \leq 1}$.

• Now consider y s.t $\|y\|_\infty \leq 1$:

$$\sum_{i=1}^{d} y_i x_i - \sum_{i=1}^{d} |x_i| \leq \sum_{i=1}^{d} |y_i||x_i| - \sum_{i=1}^{d} |x_i|$$

$$= \sum_{i=1}^{d} |x_i| \underbrace{(|y_i| - 1)}_{\leq 0} \leq 0$$

For example taking $x=0$ we get 0

} 0 is the sup

Finally, putting everything together:

$$\boxed{f^*(y) = \begin{cases} 0 & \text{if } \|y\|_\infty \leq 1 \\ +\infty & \text{otherwise} \end{cases}}$$

2) We can rewrite the problem $\min_x \|Ax - b\|_2^2 + \|x\|_1$ as

$$\min \|t\|_2^2 + \|x\|_1$$
$$\text{st } t - Ax + b = 0$$

$$L(x, t, \mu) = \|t\|_2^2 + \|x\|_1 + \mu^T(t - Ax + b)$$

The dual function is

$$g(\mu) = \inf_{x,t} \{ \|t\|_2^2 + \|x\|_1 + \mu^T(t - Ax + b) \}$$

$$= \mu^T b + \underbrace{\inf_t \{ \|t\|_2^2 + \mu^T t \}}_{(*)} - \underbrace{\sup_x \{ \mu^T A x - \|x\|_1 \}}_{= f^*(A^T \mu)}$$

(*) $\frac{\partial}{\partial t}(t^T t + \mu^T t) = 2t + \mu = 0 \Rightarrow t^* = -\frac{1}{2}\mu$

Evaluating in t^* $\quad t^T t + \mu^T t = +\frac{1}{2}\mu^T \mu (+\frac{1}{2}) + \mu^T(-\frac{1}{2}\mu)$

Note that $\frac{\partial^2}{\partial t^2}(t^T t + \mu^T t) = 2 > 0$ $\quad = \frac{1}{4}\mu^T \mu - \frac{1}{2}\mu^T \mu$

then the solution is a minimum $\quad = -\frac{1}{4}\mu^T \mu$

Convex Optimization

Ex2) 2) (cont.)

$$g(\mu) = \mu^T b - \frac{1}{4}\mu^T\mu - f^*(A^T\mu)$$

Now we want to maximize $g(\mu)$

Then we need $f^*(A^T\mu) = 0$ (otherwise we don't get the max of g)

Finally the dual problem is

$$\boxed{\begin{aligned} &\max \ \mu^T b - \frac{1}{4}\mu^T\mu \\ &\text{st} \ \|A^T\mu\|_\infty \leq 1 \end{aligned}}$$

Convex Optimization

Ex3| We want to solve $\min_{w} \frac{1}{n}\sum_{i=1}^{n} \mathcal{L}(w, x_i, y_i) + \frac{\tau}{2}\|w\|_2^2$ (sep1)

with $\mathcal{L}(w, x_i, y_i) = \max\{0, 1 - y_i(w^T x_i)\}$

1) Let $\min_{w,z} \frac{1}{n\tau} 1^T z + \frac{1}{2}\|w\|_2^2$ st $z_i \geq 1 - y_i(w^T x_i)$ $i = 1, \dots, n$ (sep2)

$z \geq 0$

$$1^T z = \sum_{i=1}^{n} z_i$$

Then (sep 2) is equivalent to

$$\min_{w,z} \frac{1}{n\tau} \sum_{i=1}^{n} \max\{0, z_i\} + \frac{1}{2}\|w\|_2^2 \quad \text{st} \quad z_i \geq 1 - y_i(w^T x_i) \ \forall i = 1, \dots, n$$

Note that in order to achieve the minimum we need the smallest value of z_i, thus $z_i = 1 - y_i(w^T x_i)$

Finally this proves that solving (sep 2) we get the solution to (sep 1)

Escaneado con CamScanner

Convex Optimization

Ex3| 2) Standard form of (sep 2) → convex problem

$$\min_{w,z} \frac{1}{n\tau} \mathbb{1}^T z + \frac{1}{2}\|w\|_2^2 \quad \text{s.t} \quad -z \leq 0$$
$$1 - y_i(w^T x_i) - z_i \leq 0 \quad i=1,\dots,n$$

$$L(w,z,\lambda,\pi) = \frac{1}{n\tau}\mathbb{1}^T z + \frac{1}{2}\|w\|_2^2 + \sum_{i=1}^n \lambda_i (1 - y_i(w^T x_i) - z_i) - \pi^T z$$

• Dual function

$$g(\lambda,\pi) = \inf_{w,z} L(w,z,\lambda,\pi)$$

$$\frac{\partial g}{\partial w} = \underbrace{w}_{\in\mathbb{R}^d} - \sum_{i=1}^n \lambda_i y_i \underbrace{x_i}_{\in\mathbb{R}^d} = 0 \rightarrow w = \sum_{i=1}^n \lambda_i y_i x_i$$

L affine on z:

$$\frac{\mathbb{1}}{n\tau} - \pi - \lambda = 0 \quad , \text{ if not } g(\lambda,\pi) = -\infty$$

$$w^T x_j = \sum_{i=1}^n \lambda_i y_i x_i^T x_j$$

$$\|w\|_2^2 = w^T w = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j y_i y_j x_i^T x_j$$

$$\sum_{i=1}^n \lambda_i (1 - y_i(w^T x_i)) = \sum_{i=1}^n \lambda_i \left[1 - y_i\left(\sum_{j=1}^n \lambda_j y_j x_j^T x_i\right)\right]$$

$$g(\lambda,\pi) = \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \lambda_i\lambda_j y_i y_j x_i^T x_j + \sum_{i=1}^n \lambda_i\left[1 - y_i\left(\sum_{j=1}^n \lambda_j y_j x_i^T x_j\right)\right]$$

$$\boxed{g(\lambda) = \begin{cases} \sum_{i=1}^n \lambda_i - \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \lambda_i\lambda_j y_i y_j x_i^T x_j & \text{if } \frac{\mathbb{1}}{n\tau} - \pi - \lambda = 0 \\ -\infty & \text{otherwise} \end{cases}}$$

• Dual problem

$$\max_{\lambda,\pi} \sum_{i=1}^n \lambda_i - \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \lambda_i\lambda_j y_i y_j x_i^T x_j$$
$$\text{st } \lambda \geq 0$$
$$\frac{\mathbb{1}}{n\tau} - \pi - \lambda = 0$$

Convex Optimization : Optionals

Ex 4) $\min \; c^T x \quad \text{s.t} \quad \sup_{a \in \mathcal{P}} (a^T x) \leq b$, with $x \in \mathbb{R}^n$

$\mathcal{P} = \{ a / C^T a \preceq d \}$

non empty polyhedra

Let's consider $\max_{a \in \mathcal{P}} a^T x$ (Hint)

In standard form this is $\min_a \; -a^T x \quad \text{st} \quad C^T a - d \preceq 0$ (*)

- Lagrangean $L(a, \lambda) = -a^T x + \lambda^T (C^T a - d) \to$ affine on a
- Dual function

$$g(\lambda) = \inf_a \left\{ -a^T x + \lambda^T (C^T a - d) \right\}$$

$$g(\lambda) = \begin{cases} -\lambda^T d & \text{if} \quad -x + C\lambda = 0 \\ -\infty & \text{otherwise} \end{cases}$$

- Dual problem

$$\max_{\substack{\lambda \succeq 0 \\ x = C\lambda}} -\lambda^T d \;=\; \min_{\substack{\lambda \succeq 0 \\ x = C\lambda}} \lambda^T d$$

Note that (*) is a convex problem + $\mathcal{P}$ non-empty by hypothesis

Then strong duality holds ($d^* = p^*$)

Because of that, we can rewrite the first problem as

$$\min_x \; c^T x \quad \text{st} \quad \min_{\substack{\lambda \succeq 0 \\ x = C\lambda}} \lambda^T d \leq b$$

which is equivalent to:

$$\boxed{\min_{x, \lambda} \; c^T x \quad \text{st.} \quad \begin{array}{l} \lambda^T d \leq b \\ \lambda \succeq 0 \\ x = C\lambda \end{array}}$$

Convex Optimization, Optionals

Ex 5

1) Lagrange relaxation

$$\min \ c^T x$$
$$\text{st} \quad Ax - b \leq 0$$
$$x_i(1-x_i) = 0 \quad \forall i = 1, \dots, n$$

- Lagrangean

$$L(x, \lambda, \mu) = c^T x + \lambda^T (Ax - b) + \sum_{i=1}^{n} \mu_i x_i (1 - x_i)$$

- Dual function

$$g(\lambda, \mu) = \inf_x \Big\{ c^T x + \lambda^T (Ax - b) + \mu^T x - \underbrace{\sum_{i=1}^{n} \mu_i x_i^2}_{= x^T M x} \Big\}$$

$$M = \begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{pmatrix}$$

$$= \inf_x \Big\{ c^T x + \lambda^T A x - \lambda^T b + \mu^T x - x^T M x \Big\}$$

- If $\mu \geq 0$ then $g(\lambda, \mu) = -\infty$
- If $\mu < 0 \rightarrow$ positive quadratic

$$\frac{\partial L}{\partial x} = c + A^T \lambda + \mu - 2Mx = 0$$

$$x = \frac{1}{2} M^{-1} (c + A^T \lambda + \mu) \qquad M^{-1} = \begin{pmatrix} \frac{1}{\mu_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\mu_n} \end{pmatrix}$$

$$g(\lambda, \mu) = -\lambda^T b + \frac{1}{2} \Big[c^T + \lambda^T A + \mu^T - \frac{1}{2} (c + A^T \lambda + \mu)^T \underbrace{M^{-1} M}_{I} \Big] M^{-1} (c + A^T \lambda + \mu)$$

$$= -\lambda^T b + \frac{1}{2} \Big[c^T + \lambda^T A + \mu^T - \frac{1}{2} (c^T + \lambda^T A + \mu) \Big] M^{-1} (c + A^T \lambda + \mu)$$

$$= -\lambda^T b + \frac{1}{4} [c^T + \lambda^T A + \mu^T] M^{-1} [c + A^T \lambda + \mu]$$

$$= -\lambda^T b + \frac{1}{4} (c + A^T \lambda + \mu)^T M^{-1} (c + A^T \lambda + \mu)$$

$$g(\lambda, \mu) = \begin{cases} -\lambda^T b + \dfrac{1}{4} \displaystyle\sum_{i=1}^{n} \dfrac{(c_i + A_i^T \lambda + \mu_i)^2}{\mu_i} & \text{if } \mu < 0 \\ -\infty & \text{otherwise} \end{cases}$$

Convex Optimization: Optionals

Ex 5 (cont.)

• Dual problem

$$\max_{\substack{\lambda,\mu \\ \lambda \geq 0, \mu \leq 0}} -\lambda^T b + \frac{1}{4}\sum_{i=1}^{n} \frac{(c_i + A_i^T\lambda + \mu_i)^2}{\mu_i}$$

using the hint

$$\max_{\substack{\lambda \\ \lambda \geq 0}} -\lambda^T b + \sum_{i=1}^{n} \min\{0, (c_i + A_i^T\lambda)\}$$

2) $\min \ c^T x$

st $Ax - b \leq 0$; $-x_i \leq 0$; $x_i - 1 \leq 0 \ \forall i = 1, \ldots, n$

• Lagrangean

$$L(x, \lambda_1, \lambda_2, \lambda_3) = c^T x + \lambda_1^T (Ax - b) - \lambda_2^T x + \lambda_3^T x - \lambda_3^T \mathbb{1}$$

• Dual function

$g(\lambda_1, \lambda_2, \lambda_3) = \inf_x L(x, \lambda_1, \lambda_2, \lambda_3)$, L is affine on x

$$g(\lambda_1, \lambda_2, \lambda_3) = \begin{cases} -\lambda_1^T b - \lambda_3^T \mathbb{1} & \text{if } c + A^T\lambda_1 - \lambda_2 + \lambda_3 = 0 \\ -\infty & \text{otherwise} \end{cases}$$

• Dual problem

$$\max_{\lambda_1,\lambda_2,\lambda_3} -\lambda_1^T b - \lambda_3^T \mathbb{1} \quad \text{st} \quad c + A^T\lambda_1 - \lambda_2 + \lambda_3 = 0$$
$$\lambda_1 \geq 0, \ \lambda_2 \geq 0, \ \lambda_3 \geq 0$$

$$\lambda_3 = \lambda_2 - c - A^T\lambda_1$$

$$(\lambda_2 - c - A^T\lambda_1)^T \mathbb{1} = \sum_{i=1}^{n} \underbrace{(\lambda_{2i} - c_i - A_i^T\lambda_1)}_{\max\{0, \lambda_{2i} - c_i - A_i^T\lambda_1\}}$$

$$\Rightarrow -\lambda_3^T \mathbb{1} = -\sum_{i=1}^{n} \max\{0, \lambda_{2i} - c_i - A_i^T\lambda_1\} \quad \text{since } \lambda_3 \geq 0$$
$$= + \sum_{i=1}^{n} \min\{0, c_i + A_i^T\lambda_1 - \lambda_{2i}\}$$

Then maximizing on λ_2 we get

$$\max_{\lambda_1 \geq 0} -\lambda_1^T b + \sum_{i=1}^{n} \min\{0, c_i + A_i^T\lambda_1\}$$

which is equivalent to the dual of (1)