Elías Masquil

Convex Optimization - Homework 2

Ex 1

$c \in \mathbb{R}^d$, $b \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times d}$

$$(P) \quad \min_x \ c^T x \quad \text{s.t } Ax = b, \ x \geq 0 \qquad\qquad (D) \quad \max_y \ b^T y \quad \text{st } A^T y \leq c$$

1) (P) in its standard form

$$\min_x \ c^T x \quad \text{st } Ax - b = 0, \ -x \leq 0$$

$$\mathcal{L}(x, \lambda, \mu) = c^T x - \lambda^T x + \mu^T (Ax - b)$$

Dual function is $g(\lambda, \mu) = \inf_x L(x, \lambda, \mu)$

Since $\mathcal{L}$ is convex in x, let's find the minimum by setting the gradient to 0

$$\nabla \mathcal{L}_x (x, \mu, \lambda) = c - \lambda + A^T \mu = 0$$

$$c = \lambda - A^T \mu$$

If $c = \lambda - A^T \mu$

$$g(\lambda, \mu) = (\lambda^T - \mu^T A) x - \lambda^T x + \mu^T (Ax - b)$$

$$g(\lambda, \mu) = -\mu^T b$$

, more generally $\boxed{g(\lambda, \mu) = \begin{cases} -\mu^T b & \text{if } c = \lambda - A^T \mu \\ \infty & \text{otherwise} \end{cases}}$

Then, the dual problem is

$$\max_{\lambda, \mu} \ -\mu^T b \quad \text{St } \lambda \geq 0 \qquad = \qquad \boxed{\max_{\lambda, \mu} \ -\mu^T b \quad \text{st } \lambda \geq 0, \ c - \lambda + A^T \mu = 0}$$

2) (D) in standard form

$$\min_y \ -b^T y \quad \text{st } A^T y - c \leq 0$$

$$\mathcal{L}(y, \lambda) = -b^T y + \lambda^T (A^T y - c)$$

$$\nabla \mathcal{L}(\lambda) = -b + A\lambda = 0 \iff b = A\lambda$$

$$\Rightarrow g(\lambda) = \begin{cases} -\lambda^T c & \text{if } b - A\lambda = 0 \\ \infty & \text{otherwise} \end{cases}$$

And the dual problem is $\boxed{\max_\lambda \ -\lambda^T c \quad \text{s.t } \lambda \geq 0, \ b - A\lambda = 0}$

3) Prove the following problem is self-dual

$$\min_{x,y} \quad c^T x - b^T y$$
$$\text{s.t} \quad Ax = b$$
$$x \geq 0$$
$$A^T y \leq c$$

Since x, y are decoupled, the problem can be solved independently for x and y

$$\min_x \quad c^T x \qquad\qquad \max_y \quad b^T y$$
$$\text{st} \quad Ax = b \qquad\qquad \text{st} \quad A^T y \leq c$$
$$x \geq 0$$

The dual problems are (from the previous part)

$$\max_{\nu,\mu} \quad -\mu^T b \qquad \text{and} \qquad \max_\lambda \quad -\lambda^T c$$
$$\text{st} \quad \nu \geq 0 \qquad\qquad \text{st} \quad \lambda \geq 0$$
$$c - \nu + A^T \mu = 0 \qquad\qquad \underbrace{b - A\lambda = 0}$$

which is equivalent as

$$\min_\lambda \quad \lambda^T c$$
$$\text{st} \quad \lambda \geq 0$$
$$b - A\lambda = 0$$

Joining both duals again:

$$\min_{\lambda,\mu} \quad \lambda^T c + \mu^T b$$
$$\text{st} \quad \nu \geq 0$$
$$\lambda \geq 0$$
$$c - \nu + A^T \mu = 0$$
$$b - A\lambda = 0$$

By observing that $x = x^T$ if $x \in \mathbb{R}$

We see that the dual of the original problem is:

$$\min_{\lambda,\mu} \quad c^T \lambda + b^T \mu \qquad\qquad \min_{\lambda,\mu} \quad c^T \lambda + b^T \mu$$
$$\text{st} \quad \nu \geq 0 \qquad\qquad \text{st} \quad c + A^T \mu \geq 0$$
$$\lambda \geq 0 \quad = \quad \lambda \geq 0$$
$$c - \nu + A^T \mu = 0 \qquad\qquad b - A\lambda = 0$$
$$b - A\lambda = 0$$

Now, doing the change of variable $\lambda = x$; $\mu = -y$

$$\boxed{\begin{array}{lll} \min\limits_{x,y} & c^T x - b^T y & \\ & \text{st} & c \geq A^T y \\ & & x \geq 0 \\ & & b = Ax \end{array}}$$

Elías Masquil

Convex Optimization - Homework 2

Ex 1

4) $\min_{x,y} \; c^T x - b^T y$ (self-Dual)

st $Ax = b$
$x \geq 0$
$A^T y \leq c$

is equivelet as minimizing over one veriable at a time, since the constraints are independet

So if x^* is optimal of Self-Dual is also optimal of P and viceversa.

$$\min_x c^T x \quad (P)$$
st $Ax = b$
$x \geq 0$

The same is valid for y^*.

From part 1)

The dual of $\min_x c^T x$
st $Ax = b$
$x \geq 0$

is $\max_{\lambda, \mu} -\mu^T b$
$\lambda \geq 0$
$c - \lambda + A^T \mu = 0$
which is eq to

$\max_{\mu} -\mu^T b$
$c + A^T \mu \geq 0$
eq. $\max_y y^T b$
$A^T y \leq c$

Finally

$$\min_{x,y} c^T x - b^T y \;\; \text{st } Ax = b,\; x \geq 0,\; A^T y \leq c \;\; = \;\; \min_x c^T x \;\; \text{st } Ax = b,\; x \geq 0 \;\; - \min_{A^T y \leq c} b^T y$$

Since there's strong duality : $c^T x^* = y^* b$

so the self-dual optimal value is $c^T x^* - b^T y^* = 0$

In summary : self DUAL = P − D and because of strong duality $P^* = D^*$ $\Rightarrow$ SELF DUAL $^* = P^* - D^* = 0$

Elías Masquil

Convex optimization - Homework 2

Ex 2

$A \in \mathbb{R}^{n \times d}$, $b \in \mathbb{R}^n$

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

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

$$f^*(y) = \sup_x \; y^T x - \|x\|_1$$

$$f^*(y) = \sup_x \; \sum_1^d y_i x_i - \sum_1^d |x_i|$$

- Supose $y / y_j > 1$ and $x / x_j = K$ $(K > 0)$

$$K y_j - K = K \underbrace{(y_j - 1)}_{\substack{>0 \\ \text{because} \\ y_j > 0}} \xrightarrow[K \to \infty]{} \infty$$

So, if $y / y_j > 1$, $y \notin \operatorname{dom} f^*(y)$

- Furthermore, if $\sum y_j > d$ and $x = \vec{1} K$

$$K \sum y_j - dK = K \underbrace{(\sum y_j - d)}_{>0} \xrightarrow[K \to \infty]{} \infty$$

So $y / \sum y_j > d \notin \operatorname{dom} f^*(y)$

- Then if $\sum y_j < -d$ $\quad x = \vec{1}(-K)$ $\quad K > 0$

$$-K \sum y_j - dK = -K \underbrace{(\sum y_j + d)}_{<0}$$

let's call it $-\beta$ $(\beta > 0)$

$$= \beta K \xrightarrow[K \to \infty]{} \infty$$

- $y / y_j < -1$ and $x / x_j = -K$ $(K > 0)$ and 0 otherwise

$y = \begin{pmatrix} \vdots \\ -\alpha \\ \vdots \end{pmatrix}$ $\alpha > 1$

$$\alpha K - |-K| = K \underbrace{(\alpha - 1)}_{\substack{\geq 0 \\ \alpha > 1}} \xrightarrow[K \to \infty]{} \infty$$

$$\sum y_i x_i - \sum |x_i| \leq \sum |y_i x_i| - \sum |x_i| = \sum (|y_i|-1)|x_i|$$

So

$$\sum y_i x_i - \sum |x_i| \leq \underbrace{\sum (|y_i| - 1)|x_i|} \leq 0$$

This is only bounded if $|y_i| \leq 1 \ \forall i$

And the bound is 0, obtained when $|y_i| = 1 \ \forall i$

Summarizing the results:

$y \in \text{dom } f^*$ if:

- $\|y\|_\infty < 1$ which takes into account all the previous results: $\sum y_i < d$, $\sum y_i > -d$, $y_i \leq 1$, $y_j \geq -1$

Now, for verifying that 0 is the supremum, let's find x / $y^T x - \sum_i |x_i| = 0$

For example $x = \vec{0}$: $\sum y_i \cdot 0 - \sum 0 = 0$

so 0 it's the supremum Then

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

2) $\min_{x,y} \|y\|_2^2 + \|x\|_1$ is equivalent to the original unconstrained problem.

st $y = Ax - b$

$$\mathcal{L}(x, y, \mu) = \|y\|_2^2 + \|x\|_1 + \mu^T (y - Ax + b)$$

The dual function is:

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

For the inf on y let's derive (note that the variables are decoupled)

$$2y + \mu = 0 \Rightarrow y = -\mu \cdot \frac{1}{2}$$

$$g(\mu) = \inf_x \frac{1}{4}\|\mu\|_2^2 - \frac{\|\mu\|_2^2}{2} + \mu^T b + \|x\|_1 - \mu^T A x$$

Elías Masquil

Convex Optimization - Homework 2

$$g(\mu) = -\sup_{x} \frac{\|\mu\|_2^2}{4} - \mu^T b + \mu^T A x - \|x\|_1$$

$$g(\mu) = -\frac{\|\mu\|_2^2}{4} + \mu^T b + f^*(A^T\mu)$$

→ from part 1

The dual problem is

$$\max_{\mu} -\frac{\|\mu\|_2^2}{4} + \mu^T b + f^*(A^T\mu)$$

$$\boxed{\begin{aligned}&\max_{\mu} -\frac{\|\mu\|_2^2}{4} + \mu^T b \\ &\quad \text{s.t. } \|A^T\mu\|_\infty \leq 1\end{aligned}}$$

because if $\|\mu^T A\| > 1$, $f^*(\mu^T A) = \infty$

Elias Masquil

Convex Optimization - Homework 2

Ex 3

$$f(w, x_i, y_i) = \max\{0 ; 1 - y_i(w^T x_i)\}$$

Sep 1 : $\min_{w} \frac{1}{n} \sum_{i}^{n} f(w, x_i, y_i) + \frac{\tau}{2} \|w\|_2^2$

1) sep 2: $\min_{w, z} \frac{1}{n\tau} \mathbb{1}^T z + \frac{1}{2} \|w\|_2^2$

$$\text{st} \quad z_i \geq 1 - y_i(w^T x_i) \quad \forall i$$
$$z \geq 0$$

Sep 2 is equivalent to the following because it still only sums over $z \geq 0$

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

Which is equivalet to this, considering $z > 0$

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

Note that the optimal z^* will be such that $z^* = 1 - y_i(w^{*T} x_i)$ $\forall i$, since all the z_i terms in the objective function are positive. If $z^* > 1 - y_i(w^{*T} x_i)$ it would be possible to take another z^{**} close to z^* that will still validate the inequality restriction, but $\sum \max\{0, z_i^{**}\} < \sum \max\{0, z_i^*\}$ contradicting the hypothesis that z^* was the optimal value.
This holds because $\forall w$ the z that minimizes the objective will be the smallest possible z, which still holds for the optimal w^*.

Since sep 1 and sep 2 have equivalent objective functions and the solution to sep 2 satisfies the constraints of sep 1, solving sep 2 is equivalent to solving sep 1.

2)

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

$$g(\lambda, \pi) = \inf_z \frac{\sum z_i}{n\tau} - \sum \pi_i z_i - \sum \lambda_i z_i + \inf_w \frac{\|w\|^2}{2} - \sum \lambda_i y_i w^T x_i$$

$$+ \sum \lambda_i$$

a) $\inf_z \frac{\sum z_i}{n\tau} - \sum \pi_i z_i - \sum \lambda_i z_i = \begin{cases} -\infty \\ 0 \end{cases}$ depending on π

To avoid $-\infty$, all terms with z_i must be 0 since it's affine in z

$$\frac{1}{n\tau} - \pi_i - \lambda_i = 0 \Leftrightarrow \boxed{\pi_i = \frac{1}{n\tau} - \lambda_i}$$

b) $\inf_w \frac{\|w\|^2}{2} - \sum \lambda_i y_i w^T x_i$

Deriving and setting the derivative to 0 (it's convex in w)

$$w - \sum_i \lambda_i y_i \vec{x_i} = 0$$

$$\boxed{w = \sum_i \lambda_i y_i x_i}$$

Plugging everything back in g:

$$g(\lambda, \pi) = \cancel{\frac{\sum z_i}{n\tau}} - \cancel{\frac{\sum z_i}{n\tau}} + \cancel{\sum z_i \lambda_i} - \cancel{\sum \lambda_i z_i}$$

$$+ \frac{1}{2} \sum_i \sum_j \lambda_i y_i \lambda_j y_j x_i^T x_j - \sum_i \lambda_i y_i \sum_j \lambda_j y_j x_j^T x_i$$

$$+ \sum \lambda_i$$

$$\boxed{g(\lambda, \pi) = \begin{cases} \sum \lambda_i - \frac{1}{2} \sum_i \sum_j \lambda_i y_i \lambda_j y_j x_i^T x_j & \text{if } \pi = \frac{1}{n\tau} - \lambda \\ -\infty & \text{otherwise} \end{cases}}$$

Elias Masquil

Convex optimization - Homework 2

Ex 4

Robust linear programming

$$\min \; c^T x$$
$$\text{st} \; \sup_{a \in P} a^T x \leq b$$

$P = \{a / C^T a \leq d\}$ (non empty polyhedra)

Show that is equivalet to

$$\min \; c^T x$$
$$\text{st} \; d^T z \leq b$$
$$C^T z = x$$
$$z \geq 0$$

Hint $\max \; a^T x$
$\text{st} \; a \in P$

In the standard form $\min \; -a^T x$
$\text{st} \; C^T a \leq d$

$$\mathcal{L}_0(a, \lambda) = -a^T x + \lambda^T (C^T a - d)$$

$$g(\lambda) = \inf_a \; -a^T x + \lambda^T (C^T a - d)$$

To make g bounded terms in a must cancel each other

$$-x + C\lambda = 0$$

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

The dual problem is

$$\max_\lambda \; -\lambda^T d \quad \text{st} \; x = C\lambda, \; \lambda \geq 0 \quad = \quad \boxed{\min_\lambda \; \lambda^T d \quad \text{st} \; x = C\lambda, \; \lambda \geq 0}$$

Note that this problem is convex and Slater's constraint holds because P is non empty by hypothesis. Then, there's strong duality, which implies that $\exists \; z /$

$$\max \; a^T x \;\; \text{st} \; a \in P \quad = \quad z^T d \;\; \text{st} \; x = Cz, \; z \geq 0$$

Finally note that the original problem can be expressed as

$$\begin{aligned} \min \quad & c^T x \\ \text{st} \quad & \begin{pmatrix} \max & a^T x \\ \text{st} & a \in P \end{pmatrix} \leq b \end{aligned}$$

By applying the previous solution we have

$$\boxed{\begin{aligned} \min \quad & c^T x \\ \text{st} \quad & z^T d \leq b \\ & z \succeq 0 \\ & Cz = x \end{aligned}}$$

Elias Masquil

Convex Optimization - Homework 2

Ex 5 Boolean LP

$$\min \ c^T x$$
$$\text{st} \quad Ax \le b$$
$$x_i \in \{0, 1\}$$

Relaxation

$$\min \ c^T x$$
$$Ax \le b$$
$$0 \le x_i \le 1$$

1) Lagrangian Relaxation

$$\min \ c^T x$$
$$\text{st} \quad Ax \le b$$
$$x_i(1-x_i) = 0$$

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

The dual function: $g(\lambda, \mu) = \inf_x c^T x + \lambda^T (Ax - b) + \underbrace{\sum_i \mu_i x_i (1 - x_i)}_{\mu^T x - x^T \operatorname{diag}(\mu) x}$

$$g(\lambda, \mu) = \inf_x c^T x + \lambda^T A x + \mu^T x - x^T \operatorname{diag}(\mu) x - \lambda^T b$$

$\mathcal{L}$ is a negative quadratic form so the inf is $-\infty$ if $\mu \geq 0$.

If $\mu < 0$ the minimum can be found by setting the derivative to 0.

$$\Rightarrow \quad c + A^T \lambda + \mu - 2 \operatorname{diag}(\mu) x = 0$$

$$x = \frac{1}{2} \operatorname{diag}\left(\frac{1}{\mu}\right) (c + A^T \lambda + \mu)$$

Plugging this in $\mathcal{L}$

$$g(\lambda, \mu) = (c^T + \lambda^T A + \mu^T) \left(\frac{1}{2} \operatorname{diag}(1/\mu)(c + A^T \lambda + \mu) \right)$$
$$- \frac{1}{2} (c^T + \lambda^T A + \mu^T) \underbrace{\operatorname{diag}(1/\mu) \operatorname{diag}(\mu)}_{Id} \frac{1}{2} \operatorname{diag}(1/\mu) (c + A^T \lambda + \mu)$$
$$- \lambda^T b$$

$$g(\lambda,\mu) = [c^T + \lambda^T A + \mu^T]\, \frac{1}{2} \operatorname{diag}(1/\mu)\, [c + A^T\lambda + \mu]$$
$$- \frac{1}{4} [c^T + \lambda^T A + \mu^T] \operatorname{diag}(1/\mu)\, [c + A^T\lambda + \mu]$$
$$- \lambda^T b$$

$$g(\lambda,\mu) = -\lambda^T b + \frac{1}{4} B \operatorname{diag}(1/\mu) B^T$$

this is an inner product weighted by $1/\mu_i$

$$\boxed{g(\lambda,\mu) = -\lambda^T b + \frac{1}{4} \sum_i \frac{1}{\mu_i} (c_i + a_i^T\lambda + \mu_i)^2 \quad \text{if } \mu < 0}$$

a_i: i-row of A

Then, the dual problem is:

$$\max_{\lambda,\mu} \; -\lambda^T b + \frac{1}{4} \sum_i \frac{1}{\mu_i} (c_i + a_i^T\lambda + \mu_i)^2$$
$$\text{st} \quad \mu < 0, \; \lambda \geq 0$$

Which is equivalent as

$$\max_{\lambda,v} \; -\lambda^T b + \frac{1}{4} \sum_i \frac{1}{-v_i} (c_i + a_i^T\lambda - v_i)^2$$
$$\text{st} \quad v > 0, \; \geq 0$$

Using the hint, minimizing over v and then over λ

The problem is equivalent to

$$\max_{\lambda} \; -\lambda^T b + \sum_i \min\{0, c_i + a_i^T\lambda\}$$
$$\lambda \geq 0$$

Elías Masquil

Convex Optimization - Homework 2

Ex 5

2) $\mathcal{L}_0(x, \lambda_1, \lambda_2, \lambda_3) = c^T x + \lambda_1^T(Ax - b) - \lambda_2^T x + \lambda_3^T(x - \vec{1})$

The dual function will be $-\infty$ if the x terms remains
Now I'm eliminating those terms

$\Rightarrow c^T + \lambda_1^T A - \lambda_2^T + \lambda_3^T = 0$

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

$$\boxed{\begin{array}{l} \text{The dual problem is} \\ \max_{\lambda_1, \lambda_2, \lambda_3} \; -\lambda_1^T b - \lambda_3^T \mathbb{1} \\ \text{st} \quad c^T + \lambda_1^T A - \lambda_2^T + \lambda_3^T = 0 \\ \qquad \lambda_1, \lambda_2, \lambda_3 \geq 0 \end{array}}$$

Using $\begin{cases} -\lambda_3^T = c^T + \lambda_1^T A - \lambda_2^T \\ \lambda_3 \geq 0 \\ \lambda_3^T \mathbb{1} = \sum \lambda_{3i} \end{cases}$

$$\max_{\lambda_1, \lambda_2} \; -\lambda_1^T b + \sum_i \min\{(c_i + \lambda_1^T a_i - \lambda_{2i}), 0\}$$
$$\text{st} \quad \lambda_1, \lambda_2 \geq 0$$

Maximizing first over λ_2, results in $\lambda_2 = 0$
because λ_2 must be ≥ 0 and is multiplied by (-1).

Then the problem is equivalent to

$$\boxed{\begin{array}{l} \max_{\lambda_1} \; -\lambda_1^T b + \sum_i \min\{0, c_i + \lambda_1^T a_i\} \\ \text{st} \quad \lambda_1 \geq 0 \end{array}}$$

Which is the same dual problem as in part 1.