DI PIAZZA
Théo

HW2 - Convex Optimization

Exercice 1

* $c \in \mathbb{R}^d$, $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times d}$

* $$\min_x c^T x \quad \text{s.t. } Ax = b,\ x \geq 0 \qquad (P)$$

* $$\max_y b^T y \quad \text{s.t. } A^T y \leq c \qquad (D)$$

1) To start, let's write (P) with the standard form:

$$\min_x c^T x \quad \text{s.t.} \quad Ax - b = 0, \quad -x \leq 0 \qquad (P)$$

Then we define: $L(x, \lambda, \mu) = c^T x + \mu^T (Ax - b) - \lambda^T x$ with $\mu \in \mathbb{R}^m$, $\lambda \in \mathbb{R}^d$

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

Then the dual of (P) defined such that:

$$\max_{\mu, \lambda} \inf_x c^T x + \mu^T (Ax - b) - \lambda^T x \quad \text{s.t. } \lambda \geq 0 \qquad (P^*)$$

Two cases are possible:

- If $(c^T + \mu^T A - \lambda^T) = 0$ then $g(\lambda, \mu) = -\mu^T b$
- Else $g(\lambda, \mu) = -\infty$

Then we can rewrite (P^*) such that:

$$\max_{\mu, \lambda} -\mu^T b \quad \text{s.t. } \mu^T A + c^T - \lambda^T = 0, \quad \lambda \geq 0$$

which can be simplified:

$$\max_{\mu} -\mu^T b \quad \text{s.t. } \mu^T A + c^T \geq 0$$

2) Let's write (D) with standart form:

$$\min_{y} \; -b^T y \qquad (D)$$
$$\text{s.t. } A^T y - c \leq 0$$

Then we can define: $L(y,\lambda) := -b^T y + \lambda^T (A^T y - c)$ with $\lambda \in \mathbb{R}^m$

$$g(\lambda) := \inf_{y} -b^T y + \lambda^T (A^T y - c)$$

Then the dual of (D), noted (D^*) can be defined such that:

$$\max_{\lambda} \inf_{y} -b^T y + \lambda^T (A^T y - c)$$
$$\text{s.t. } \lambda \geq 0$$

Two cases are possible:

- If $-b^T + \lambda^T A^T = 0$ then $g(\lambda) = -\lambda^T c$
- Else, $g(\lambda) = -\infty$

Then we can rewrite (D^*):

$$\max_{\lambda} -\lambda^T c$$
$$\text{s.t. } A\lambda = b$$
$$\lambda \geq 0$$

3)

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

Let's define: $L(x,y,\mu,\lambda_1,\lambda_2) := c^T x - b^T y + \mu^T(b - Ax) - \lambda_1^T x + \lambda_2^T (A^T y - c)$

$g(\mu,\lambda_1,\lambda_2) := \inf_{x,y} \left(c^T - \mu^T A - \lambda_1^T \right) x + \left(-b^T + \lambda_2^T A^T \right) y + \mu^T b - \lambda_2^T c$

Then, the dual of (SD), noted (SD^*) is defined:

$$\max_{\mu,\lambda_1,\lambda_2} \inf_{x,y} L(x,y,\mu,\lambda_1,\lambda_2) \qquad (SD^*)$$
$$\text{s.t.} \quad \lambda_1, \lambda_2 \geq 0$$

Two cases are possible:

- If $c^T - \mu^T A - \lambda_1^T = -b^T + \lambda_2^T A^T = 0$, then $g(\mu,\lambda_1,\lambda_2) = \mu^T b - \lambda_2^T c$
- Else, $g(\mu,\lambda_1,\lambda_2) = -\infty$

Then, (SD^*) can be simplified:

$$\max_{\mu,\lambda_1,\lambda_2} \quad \mu^T b - \lambda_2^T c$$
$$\text{s.t.} \quad A^T \mu \leq c$$
$$A \lambda_2 = b$$
$$\lambda_2 \geq 0$$

With $(\lambda_2 = x)$ and $(\mu = y)$, it is equivalent to:

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

which is exactly (SD).

$\Rightarrow$ Hence (SD) is self-dual.

4) • The objective function of (SD) is a sum of linear functions with respect to x and y.
Each constraint of the problem concerns only one variable at a time

Hence, we can rewrite (SD):

$$\underbrace{\begin{array}{ll}\min\limits_{x,y} & c^T x - b^T y \\ \text{s.t.} & Ax = b \\ & x \geq 0 \\ & A^T y \leq c\end{array}}_{(SD)} = \underbrace{\begin{array}{ll}\min\limits_{x} & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0\end{array}}_{(P)} + \underbrace{\begin{array}{ll}\min\limits_{y} & -b^T y \\ \text{s.t.} & A^T y \leq c\end{array}}_{(D)}$$

Then x^* can be obtained by solving (P)
and y^* can be obtained by solving (D)
Finally, $[x^*, y^*]$ will be an optimal solution for (SD) obtained by solving (P) and (D).

• To start, we can rewrite (SD):

$$\begin{array}{ll}\min\limits_{x,y} & c^T x - b^T y \\ \text{s.t.} & Ax = b \\ & x \geq 0 \\ & A^T y \leq c\end{array} = \begin{array}{ll}\min\limits_{x} & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0\end{array} - \left(+ \begin{array}{ll}\max\limits_{y} & b^T y \\ \text{s.t.} & A^T y \leq c\end{array} \right)$$

Then, by making a change of variable such as $z = -y$.

$$\begin{array}{ll}\min & c^T x - b^T y \\ \text{s.t.} & Ax = b \\ & x \geq 0 \\ & A^T y \leq c\end{array} = \underbrace{\begin{array}{ll}\min & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0\end{array}}_{(P)} - \left(+ \underbrace{\begin{array}{ll}\max\limits_{z} & -b^T z \\ \text{s.t.} & -A^T z \leq c\end{array}}_{(P^*)} \right)$$

* Constraints of (P) are a linear equality and a linear inequality.
* (P) is supposed feasible
* $\mathrm{dom}(c^T x) = \mathbb{R}^d$, which is open

Then by Slater's condition, if we define:

→ p^* as the solution of (P)

→ d^* as the solution of (P^*)

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

So finally,
$$\begin{aligned} \min\ & c^T x - b^T y \\ \text{s.t. } & Ax = b \\ & x \geq 0 \\ & A^T y \leq c \end{aligned} \quad = p^* - d^* = 0$$

DI FIA2ZA

Théo HW2 - Convex Optimization

Exercice 2

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

$$\min_x \|Ax - b\|_2^2 + \|x\|_1 \qquad (RLS)$$

1) Let $x, y \in \mathbb{R}^d$. We define $f(x) := \|x\|_1$

Then, $f^*(y) = \sup_{x \in \mathrm{dom} f} y^T x - f(x) = \sup_{x \in \mathrm{dom} f} \sum_{i=1}^d (y_i x_i - |x_i|)$

Three cases are possible:

* $\exists j \in [\![1, d]\!]$ s.t. $y_j > 1$ and we select x s.t. $\begin{cases} x_j = \alpha > 0 \\ x_i = 0 \quad \forall i \neq j \end{cases}$

Then $\sum_{i=1}^d y_i x_i - |x_i| = y_j \alpha - \alpha = \underbrace{\alpha}_{>0} \underbrace{(y_j - 1)}_{>0} \xrightarrow[\alpha \to +\infty]{} +\infty$

* $\exists j \in [\![1, d]\!]$ s.t. $y_j < -1$ and we select x s.t. $\begin{cases} x_j = \alpha < 0 \\ x_i = 0 \quad \forall i \neq j \end{cases}$

Then $\sum_{i=1}^d y_i x_i - |x_i| = y_j \alpha + \alpha = \underbrace{\alpha}_{<0} \underbrace{(y_j + 1)}_{<0} \xrightarrow[\alpha \to -\infty]{} +\infty$

Then from 2 previous cases, $f^*(y) = +\infty$ if $\|y\|_\infty > 1$

* Finally, if $\|y\|_\infty \leq 1$

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

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

The equality holds if $x = \vec{0}$, then $f^*(y) = 0$ if $\|y\|_\infty \leq 1$.

$\Rightarrow$ So finally: $f^*(y) = \begin{cases} +\infty & \text{if } \|y\|_\infty > 1 \\ 0 & \text{otherwise} \end{cases}$

2) Let's rewrite the problem (RLS). Let $y \in \mathbb{R}^m$.

$$\min_{x} \|y\|_2^2 + \|x\|_1$$

$$\text{s.t.} \quad y + b - Ax = 0$$

Then we can define $L(\mu, x, y) := \|y\|_2^2 + \|x\|_1 + \mu^T(y + b - Ax)$ with $\mu \in \mathbb{R}^m$

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

$$= \inf_{y} \|y\|_2^2 + \mu^T y + \inf_{x} \|x\|_1 - \mu^T A x \quad + \mu^T b$$

$$= \inf_{y} k(y) \quad - \quad \sup_{x} (A^T\mu)^T x - \|x\|_1 \quad + \mu^T b$$

* where $k(y) := \|y\|_2^2 + \mu^T y$ is convex and 1-differentiable

Hence $\nabla_y k(y) = 2y + \mu = 0 \iff y = -\frac{\mu}{2}$

$\Rightarrow$ The minimum of k is $\frac{1}{4}\|\mu\|_2^2 - \frac{1}{2}\|\mu\|_2^2 = -\frac{1}{4}\|\mu\|_2^2$

* With respect to 1), we can write that:

$$\sup_{x} (A^T\mu)^T x - \|x\|_1 = f^*(A^T\mu) = \begin{cases} +\infty & \text{if } \|A^T\mu\|_\infty > 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\Rightarrow g(\mu) = \mu^T b - \frac{1}{4}\|\mu\|_2^2 - f^*(A^T\mu)$$

So finally, the dual of (RLS) can be written:

$$\max_{\mu} \mu^T b - \frac{1}{4}\|\mu\|_2^2$$

$$\text{s.t.} \quad \|A^T\mu\|_\infty \leq 1$$

DI PIAZZA
Théo

HW2 - Convex Optimization

Exercice 3

$$\min_{w} \frac{1}{m}\sum_{i=1}^{m} \mathcal{L}(w, x_i, y_i) + \frac{\tau}{2}\|w\|_2^2 \qquad (\text{Sep } 1)$$

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

1)
$$\min_{w,z} \frac{1}{m\tau} \mathbf{1}^T z + \frac{1}{2}\|w\|_2^2$$
$$\text{s.t. } z_i \geq 1 - y_i(w^T x_i) \quad \forall i \in [\![1, m]\!] \qquad (\text{Sep } 2)$$
$$z \geq 0$$

- To start, τ is the regularization parameter hence $\tau > 0$.
We can rewrite (Sep 2) such as:

$$\min_{z} \frac{1}{m}\sum_{i=1}^{m} z_i + \frac{\tau}{2}\min_{z}\|w\|_2^2 \qquad (\text{Sep } 2)$$
$$\text{s.t. } z_i \geq 1 - y_i(w^T x_i) \quad \forall i \in [\![1, m]\!]$$
$$z \geq 0$$

- Then, it can be showned that:

$$\begin{array}{l} \min_{z} \frac{1}{m}\sum_{i=1}^{m} z_i \\ \text{s.t. } z_i \geq 1 - y_i(w^T x_i) \ \forall i \in [\![1, m]\!] \\ \quad z \geq 0 \end{array} = \min_{w} \frac{1}{m}\sum_{i=1}^{m} \max\{0, 1 - y_i(w^T x_i)\}$$

- So finally, we can write (Sep 2) such that:

$$\min_{w} \frac{1}{m}\sum_{i=1}^{m} \mathcal{L}(w, x_i, y_i) + \frac{\tau}{2}\|w\|_2^2 \quad \text{which is (Sep 1)}$$

2) To start, we can write (Sep2) with standard form:

$$\min_{w,\zeta} \frac{1}{m\tau} 1^T \zeta + \frac{1}{2}\|w\|_2^2 \qquad (\text{Sep } 2)$$

$$\text{s.t.} \quad 1 - \zeta_i - y_i(w^T x_i) \leq 0 \quad \forall i = 1..m \quad (\lambda_i)$$

$$-\zeta \preceq 0 \quad (\pi)$$

* Then, we can define: $(\lambda_i \in \mathbb{R} \; \forall i = 1..m, \; \pi \in \mathbb{R}^m)$

$$L(w,\zeta,\lambda,\pi) := \frac{1}{m\tau} 1^T \zeta + \frac{1}{2}\|w\|_2^2 - \pi^T \zeta + \sum_{i=1}^m \lambda_i \left(1 - \zeta_i - y_i(w^T x_i)\right)$$

$$g(\lambda,\pi) := \inf_{w,\zeta} \left(\frac{1}{m\tau} 1^T - \pi^T - \lambda^T\right)\zeta + \frac{1}{2}\|w\|_2^2 - \sum_{i=1}^m \lambda_i y_i (w^T x_i) + \lambda^T 1$$

* Then $g(\lambda,\pi) = \begin{cases} \inf_w \frac{1}{2}\|w\|_2^2 - \sum_{i=1}^m \lambda_i y_i (w^T x_i) + \lambda^T 1 & \text{if } \frac{1}{m\tau} 1 - \pi - \lambda = 0 \\ -\infty & \text{otherwise.} \end{cases}$

* $f(w) := \frac{1}{2}\|w\|_2^2 - \sum_{i=1}^m \lambda_i y_i (w^T x_i) + \lambda^T 1$ is convex and 1-differentiable

Hence $\nabla_w f(w) = w - \sum_{i=1}^m \lambda_i y_i x_i = 0 \implies w = \sum_{i=1}^m \lambda_i y_i x_i$

Hence $\inf_w f(w) = \frac{1}{2}\left(\sum_{i=1}^m \lambda_i y_i x_i\right)^T \left(\sum_{i=1}^m \lambda_i y_i x_i\right) - \sum_{i=1}^m \lambda_i y_i \left(\sum_{j=1}^m \lambda_j y_j x_j\right)^T x_i + \lambda^T 1$

$$= -\frac{1}{2} \sum_{1 \leq i,j \leq m} \lambda_i \lambda_j y_i y_j x_i^T x_j + \lambda^T 1$$

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

Hence, the dual of (Sep2) is:

$$\max_{\lambda,\pi} \ \lambda^T \mathbb{1} - \frac{1}{2} \sum_{1 \leq i,j \leq m} \lambda_i \lambda_j y_i y_j x_i^T x_j$$

$$\text{s.t.} \quad \lambda \geqslant 0$$

$$\pi \geqslant 0$$

$$\frac{1}{m\tau} \mathbb{1} - \pi - \lambda = 0$$

which can be simplified:

$$\max_{\lambda} \ \lambda^T \mathbb{1} - \frac{1}{2} \sum_{1 \leq i,j \leq m} \lambda_i \lambda_j y_i y_j x_i^T x_j$$

$$\text{s.t.} \quad \lambda \geqslant 0$$

$$\frac{1}{m\tau} \mathbb{1} - \lambda \geqslant 0$$