Elif NEBIOGLU

Convex Optimization

HOMEWORK 2

Exercise 1 (LP Duality)

1. The Lagrangian of problem P is:

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

which is an affine function of x, it follows that the dual function is

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

The dual problem is to maximize $g(\lambda, \nu)$ subject to $\lambda \geq 0$ and after making the implicit constraints explicit we obtain:

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

2. The standard form of (D): $\begin{cases} \min_y & -b^T y \\ \text{s.t} & A^T y - c \end{cases}$

The Lagrangian function:

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

Let suppose $D = \{ y \in \mathbb{R}^n \mid A^T y - c \leq 0 \}$

The dual Lagrangian function:

$$g(\lambda) = \inf_{y \in D} \{ L(y, \lambda) \} = \inf_{y \in D} \{ (A\lambda - b)^T y - c^T \lambda \}$$

This follows the dual function:

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

Then the dual problem is to maximize $g(\lambda)$ subject to $\lambda \geq 0$:

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

3. Prove that the following problem is self-dual:

$$\min_{x,y} \quad c^Tx - b^Ty$$

st

$$Ax = b$$
$$x \geq 0$$
$$A^Ty \leq c$$

→ The standert form of the problem:

$$\min_{x,y} \quad c^Tx - b^Ty$$

st

$$Ax - b = 0$$
$$-x \leq 0$$
$$A^Ty - c \leq 0$$

The Lagrangian is:

$$L(x,y,\lambda_1,\lambda_2,v) = c^Tx - b^Ty + v^T(b - Ax) - \lambda_1^T x + \lambda_2^T(A^Ty - c)$$
$$= b^Tv - c^T\lambda_2 + (c - A^Tv - \lambda_1)^T x + (A\lambda_2 - b)^T y$$

which is affine in x and y, the dual function is hence:

$$g(\lambda_1, \lambda_2, v) = \begin{cases} b^Tv - c^T\lambda_2 & \text{if } c - A^Tv - \lambda_1 = 0 \text{ and } A\lambda_2 - b = 0 \\ -\infty & \text{otherwise} \end{cases}$$

Therefore, the dual problem after making the implicit constraints explicit we obtain:

$$\max_{\lambda_1, \lambda_2, v} \quad -c^T\lambda_2 + b^T v$$

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

$$\Updownarrow$$

$$\max_{\lambda_1, \lambda_2, v} \quad -c^T\lambda_2 + b^T v$$

$$\text{s.t} \quad c \geq A^T v$$
$$A\lambda_2 = b$$
$$\lambda_2 \geq 0$$

$$\Updownarrow$$

$$\min_{\lambda_2, v} \quad c^T\lambda_2 - b^T v$$

$$c \geq A^T v$$
$$A\lambda_2 = b$$
$$\lambda_2 \geq 0$$

We can say that this problem is self-dual. We changed λ_2 to x and v to y (and the max into v min).

4. • We know that the constraint of (self-dual) problem is disjoint it can be written as:

$$\{(x, y) \mid Ax = b;\ x \geq 0;\ A^T y \leq c\}$$

Hence, the (self-dual) problem can be decomposed into two problems:

$$\begin{array}{l} \min \; c^T x - b^T y \\ \text{s.t} \\ Ax = b \\ x \geq 0 \\ A^T y \leq c \end{array} \iff \begin{array}{l} \min_x \; c^T x \\ \text{s.t} \\ Ax = b \\ x \geq 0 \end{array} + \begin{array}{l} \min_y \; -b^T y \\ \text{s.t} \\ A^T y \leq c \end{array}$$

$$\iff \begin{array}{l} \min_x \; c^T x \\ \text{s.t} \\ Ax = b \\ x \geq 0 \end{array} + \begin{array}{l} \max_y \; b^T y \\ \text{s.t} \\ A^T y \leq c \end{array}$$

$$\iff (P) + (D)$$

We can observe that (x^*, y^*) is a solution of the (self-dual) problem, in this manner, we can say that x^* is an optimal solution for (P) and y^* is an optimal solution for (D).

- The dual of problem P is problem D, and similarly, the dual of D corresponds back to P, although there may be a change of variables involved. Problem P is convex, feasible and bounded (according to the hypothesis), which ensures that strong duality holds. As a result (let p^* (resp. d^*) be the optimal value of P (resp. D))

$$\min_{x,y} \; c^T x - b^T y = p^* - d^* = p^* - p^* = 0$$

$$\begin{array}{l} Ax = b \\ x \geq 0 \\ A^T y \leq c \end{array}$$

Exercise 2 (Regularized least-square)

1. The cojugate fonction of $\|x\|_1$:

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

$$= \sup_x \{ y^T x - \sum_{i=1}^d |x_i| \}$$

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

$\star$ if $y_i > 1$: We choose $x_i = t > 0$

$x_k = 0 \ \forall k \neq i$

$$y^T x - \|x\|_1 = y_i t - t = t(y_i - 1) \underset{t \to \infty}{>0}$$

$$\sup_x \{ y^T x - \|x\|_1 \} = +\infty$$

$\star$ if $y_i < -1$: We choose $x_i = t < 0$

$x_k = 0 \ \forall k \neq i$

$$y^T x - \|x\|_1 = y_i t - (-t) = y_i t + t$$

$$= t(y_i + 1) \underset{t \to -\infty}{<0}$$

$$\sup_x \{ y^T x - \|x\|_1 \} = +\infty$$

$\star$ if $|y_i| < 1$

$x_i y_i \leq |x_i y_i| \leq |x_i|$

$$y^T x - f(x) \leq \sum_{i=1}^d |y_i x_i| - \sum_{i=1}^d |x_i|$$

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

$$\leq 0$$

$$\sup \{ y^T x - \|x\|_1 \} = 0$$

Conclusion

$$\|\cdot\|_1^*(y) = \begin{cases} 0 & \text{if } \|y\|_\infty \leq 1 \\ +\infty & \text{otherwise} \end{cases}$$

2. (RLS) $\Leftrightarrow$ $\begin{cases} \min_{x,y} & \|y\|_2^2 + \|x\|_1 \\ \text{st} & \\ & Ax - b = y \end{cases}$

The Lagrangian is:

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

The dual Lagrangian is:

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

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

$$= \inf_{x \in D} \{\inf_y \{\|x\|_1 + \vartheta^T Ax + y^T y - \vartheta^T y - \vartheta^T b\}\}$$

$$= \inf_x \{\|x\|_1 + \vartheta^T Ax\} + \inf_y \{y^T y - \vartheta^T y\} - b^T \vartheta$$

a) $f'(y) = \|y\|^2 + \vartheta^T y$ is convex and differentiable then we have:

$$\nabla f'(y) = 2y + \vartheta = 0$$

$$y = \vartheta/2$$

$$\inf_y \{y^T y - \vartheta^T y\} = \frac{\vartheta^T\vartheta}{4} - \frac{\vartheta^T\vartheta}{2} = -\vartheta^T\vartheta/4$$

b) $\sup_x \{ (-A^T v)^T x - \|x\|_1 \} = \begin{cases} 0 & \text{if } \|A^T v\|_\infty \leq 1 \\ +\infty & \text{otherwise} \end{cases}$

$\Rightarrow \inf_x \{ \|x\|_1 + v^T A x \} = \begin{cases} 0 & \text{if } \|A^T v\|_\infty \leq 1 \\ -\infty & \text{otherwise} \end{cases}$

(by definition 1.4)

$\therefore \quad g(v) = \begin{cases} b^T v - 1/4\, v^T v & \text{if } \|A^T v\|_\infty \leq 1 \\ -\infty & \text{otherwise} \end{cases}$

The dual problem is:

$$\max_v \; b^T v - 1/4 \|v\|_2^2$$

st

$$\|A^T v\|_\infty \leq 1$$

Exercise 3 (Data Seperation)

1. $\min_w \frac{1}{m} \sum_{i=1}^{n} \mathcal{L}(w; x_i; y_i) + \tau/2 \|w\|_2^2$: sep 1

$\Leftrightarrow \quad \min_w 1/m \sum_{i=1}^{n} \max\{0, 1 - y_i(w^T x_i)\} + \tau/2 \|w\|_2^2$

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

if $z_i = 0 \Rightarrow 1 - y_i(w^T x_i) \leq 0 = z_i$

if $z_i = 1 \Rightarrow 1 - y_i(w^T x_i) = 1 = z_i$

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

$\Leftrightarrow \quad 1 - y_i(w^T x_i) \leq z_i$

$z_i \geq 0$

$$\Rightarrow \min_{w} \; 1/n \sum_{i=1}^{n} f(w, x_i, y_i) + \tau/2 \, \|w\|_2^2$$

$$\Leftrightarrow \min_{w, z_i} \; 1/n \sum_{i=1}^{n} z_i + \tau/2 \, \|w\|_2^2$$

s.t

$$\begin{cases} 1 - y_i(w^T x_i) \leq z_i \\ z_i \geq 0 \end{cases}$$

$$\Updownarrow$$

$$\min_{w, z} \; 1/n\tau \; 1^T + 1/2 \, \|w\|_2^2 \quad (\text{sep } 2)$$

$$\begin{cases} 1 - y_i(w^T x_i) \leq z_i & \forall i = 1, \ldots, n \\ z \geq 0 & \Rightarrow (\text{sep } 2) \text{ solves problem (sep 1)} \end{cases}$$

2. The Lagrangian of (sep 2) is:

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

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

$$= 1/2 \, \|w\|^2 + (1/n\tau \; 1 - \pi - \lambda)^T z + 1^T \lambda - \sum_{i=1}^{n} \lambda_i y_i (w^T x_i)$$

The dual finction:

$$g(\lambda,\pi) = \inf_{z,w} \left(1/2 \|w\|_2^2 + (1/n\tau \, 1 - \pi - \lambda)^T z + \right.$$

$$\left. 1^T\lambda - \sum_{i=1}^{n} \lambda_i y_i (w^T x_i) \right)$$

a) $\inf_{w} g'(w) = \inf_{w} \left(1/2 \|w\|_2^2 - \sum_{i=1}^{n} \lambda_i y_i (w^T x_i) \right)$

$$\nabla g'(w) = w - \sum_{i=1}^{n} \lambda_i y_i x_i \ldots$$

($g'(w)$ is convex and differentiable)

Therefore, the minimum value is

$$+ w = \sum_{i=1}^{n} \lambda_i y_i x_i \quad (*)$$

b) $\inf_{z} g''(z) = \inf_{z} (1/n\tau \, 1 - \pi - \lambda)^T z$

$$= \begin{cases} 0 & \text{if } \quad 1/n\tau \, 1 - \pi - \lambda = 0 \\ -\infty & \text{oth.} \end{cases}$$

Put $(*)$ in $g'(w)$:

$$1/2 \left\| \sum_{1}^{n} \lambda_i y_i x_i \right\|_2^2 - \sum_{1}^{n} \lambda_i y_i \left(\sum \lambda_j y_j x_j^T x_i \right)$$

$$= 1/2 \left(\sum_i \lambda_i y_i x_i \right)^T \left(\sum_j \lambda_j y_j x_j \right) - \sum_i \sum_j \lambda_i y_i \lambda_j y_j x_j^T x_i$$

$$= 1/2 \sum_{i,j}^{n} \lambda_i y_i x_i^T \lambda_j y_j x_j - \sum_{ij} \lambda_i y_i \lambda_j y_j x_j^T x_i$$

$$= -1/2 \sum_{ij} \lambda_i \lambda_j y_i y_j x_i^T x_j$$

$$g(\lambda,\pi) = \begin{cases} 1^T\lambda - 1/2 \sum_{ij} \lambda_i \lambda_j y_i y_j x_i^T x_j & \text{if } 1/n\tau \, 1 - \pi - \lambda = 0 \\ -\infty & \text{oth.} \end{cases}$$

It follows that the dual problem is to max $g(\lambda,\pi)$ subject to $\lambda \geq 0$ and $\pi \geq 0$:

$$\max_{\lambda,\pi} \quad 1^T\lambda - 1/2 \sum_{ij} \lambda_i \lambda_j y_i y_j \, x_i^T x_j$$

$$\text{s t} \quad 1/n\tau \geq \lambda \geq 0.$$