Francisco Javier Sáez Maldonado
José Antonio Álvarez Ocete

Ejercicios Programación Lineal
Optimización

March 21, 2022

Exercise 1
Show that if S is an open set, its complement S^c is closed, and viceversa.

$\boxed{\Rightarrow}$

Let S be an open set. We know that S^c is a closed set if and only if for any sequence of elements $\{x_n\} \subset S^c$ such that $\{x_n\} \longrightarrow x$, then $x \in S^c$.

Let us use this characterization to prove that S^c is closed. Let $\{x_n\} \subset S^c$ such that $\{x_n\} \longrightarrow x$. Suppose that x is not in S^c. Then x must be in its complementary, S. Since S is an open set we know that $\exists \epsilon > 0$ such that $B(x, \epsilon) \subset S$.

Since $\{x_n\} \longrightarrow x$, for any $\delta > 0 \exists n \in \mathbb{N}$ such that $\| x - x_n \| < \delta$. In particular, for $\delta = \epsilon$ there is a element of the succesion x_n in $B(x, \epsilon) \subset S$, but $\{x_n\} \subset S^c$. This contradictions implies that x is, in fact, in S^c.

$\boxed{\Leftarrow}$

Let S be a set such that its complement S^c is closed (that is, $S^c = cl(S^c)$). Let us show that S is open.

Let $x \in S$. Then, $x \notin S^c$, which implies $x \notin cl(S^c)$. This implies that

$$\exists \epsilon > 0 \text{ such that } B(x, \epsilon) \cap S^c = \emptyset \implies B(x, \epsilon) \subset S$$

Thus, showing that S is open.

Exercise 2
If S_1, S_2 are convex subsets, prove that the following are also convex sets:

$$S_1 \cap S_2 = \{x \ : \ x \in S_1 \text{ and } x \in S_2\}$$
$$S_1 + S_2 = \{x + x' \ : \ x \in S_1, x' \in S_2\}$$
$$S_1 - S_2 = \{x - x' \ : \ x \in S_1, x0 \in S_2\}$$

- Consider $S_1 \cap S_2 = \{x \ : \ x \in S_1 \text{ and } x \in S_2\}$. Let $x, x' \in S_1 \cap S_2$ and consider the segment

  $$z \equiv \lambda x + (1 - \lambda)x', lambda \in [0, 1]$$

  Since x is in in S_1, which is a convex set, z will also be in S_1. Likewise, z will be in S_2, and thus in the intersection $S_1 \cap S_2$. Since the segment is contained in the intersection, $S_1 \cap S_2$ is a convex subset.

- Consider $S_1 + S_2 = \{x + z \ : \ x \in S_1, z \in S_2\}$. Now, let $x + z \in S_1 + S_2$ and $x' + z' \in S_1 + S_2$. Also, consider the segment

$$\begin{aligned}\lambda(x+z) + (1-\lambda)(x'+z') &= \lambda x + \lambda z + (1-\lambda)x' + (1-\lambda z') \\ &= \underbrace{\lambda x + (1-\lambda)x'}_{x'' \in S_1} + \underbrace{\lambda z + (1-\lambda)z'}_{z'' \in S_2}\end{aligned}$$

  Where the underbraces are true due to the convexity of S_1 and S_2, so we have

$$x'' + z'' \in S_1 + S_2$$

  so the segment is in the sum set, and thus, the set $S_1 + S_2$ is convex.

- Consider $S_1 - S_2 = \{x - z \ : \ x \in S_1, z \in S_2\}$. Now, let $x - z \in S_1 - S_2$ and $x' - z' \in S_1 - S_2$. Also, consider the segment

$$\begin{aligned}\lambda(x-z) + (1-\lambda)(x'-z') &= \lambda x - \lambda z + (1-\lambda)x' - (1-\lambda z') \\ &= \underbrace{\lambda x + (1-\lambda)x'}_{x'' \in S_1} - \underbrace{\left(\lambda z + (1-\lambda)z'\right)}_{z'' \in S_2}\end{aligned}$$

  Where the underbraces are true due to the convexity of S_1 and S_2, so we have

$$x'' - z'' \in S_1 - S_2$$

  so the segment is in the difference set, and thus, the set $S_1 - S_2$ is convex.

Exercise 3
If $f : S \to \mathbb{R}$ is a convex function on the convex set S, the set $S_{min} = \{x : \ x \text{ is a minimum of f}\}$ is a convex set.

Let y be the minimum of $f(x)$: $y = \min_x f(x)$. Then $S_{min} = \{x \in S : \ f(x) = y\}$. We need need to show that for all $x, x' \in S_{min}$ and for all $\lambda \in [0,1]$:

$$z \equiv \lambda x + (1-\lambda)x' \in S_{min} \leftrightarrow f(z) = y$$

Since S is convex, z is in S, and since f is also convex:

$$\begin{aligned}f(z) &= f\left(\lambda x + (1-\lambda)x'\right) \\ &\leq \lambda f(x) + (1-\lambda)f(x') \\ &= \lambda y + (1-\lambda)y \\ &= y\end{aligned}$$

where we used that $x, x' \in S_{min}$. But since y is the minimum of f, the equality holds $f(z) = y$. This means that z is in S_{min}, therefore S_{min} is a convex set.

Exercise 4
Given a quadratic form $q(w) = w^T Qw + bw + c$, with Q a symmetric $d \times d$ matrix, w, b being $d \times 1$ vectors and c a real number, derive its gradient and Hessian.

Let us start by unrolling the quadratic form expression:

$$q(w) = \sum_{i,j=1}^{d} Q_{ij} w_i w_j + \sum_{i=1}^{d} b_i w_i + c,$$

and compute the partial derivative over the $k-th$ component:

$$\frac{\partial q}{\partial w_k}(w) = \sum_{i=1}^{d} Q_{ik} w_i + \sum_{j=1}^{d} Q_{kj} w_j + b_k$$

where $k \in \{1, \ldots, d\}$. By using that Q is symmetric we obtain:

$$\frac{\partial q}{\partial w_k}(w) = 2 \sum_{j=1}^{d} Q_{kj} w_j + b_k. \tag{1}$$

That is, we are multiplying the $k-th$ row of the Q matrix and multiplying it by w. We can obtain gradient as a product of matrices using the previous expression:

$$\nabla q(w) = \begin{pmatrix} \frac{\partial q}{\partial w_1}(w) \\ \vdots \\ \frac{\partial q}{\partial w_d}(w) \end{pmatrix} = \begin{pmatrix} 2 \sum_{j=1}^{d} Q_{1j} w_j + b_1 \\ \vdots \\ 2 \sum_{j=1}^{d} Q_{dj} w_j + b_d \end{pmatrix} = 2 \begin{pmatrix} \sum_{j=1}^{d} Q_{1j} w_j \\ \vdots \\ \sum_{j=1}^{d} Q_{dj} w_j \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_d \end{pmatrix} = 2\, Qw + b$$

In order to obtain the gradient take partial derivative over the $l-th$ component in 1:

$$\begin{aligned} \frac{\partial^2 q}{\partial w_k \partial w_l}(w) &= \frac{\partial q}{\partial w_l}\left(\frac{\partial q}{\partial w_k}\right)(w) \\ &= \frac{\partial q}{\partial w_l}\left(2 \sum_{j=1}^{d} Q_{kj} w_j + b_k\right)(w) \\ &= 2\, Q_{kl} \end{aligned}$$

Hence, the Hessian matrix of q will have $2Q_{kl}$ in position (k, l). That is:

$$\text{Hess}\, q(w) = 2Q$$

Exercise 5
If $(p_1, \ldots, p_n)$ is a probability distribution, prove that its entropy $H(p_1, \ldots, p_n) =$

$-\sum_{i=1}^n p_i \log p_i$ is a concave function. Show also that its maximum is $\log n$, attained when $p_i = \frac{1}{n}$ for all i.

Exercise 6
We want to solve the following constrained restriction problem:

$$\begin{aligned} \min \quad & x^2 + 2y^2 + 4xy \\ \text{s.t} \quad & x + y = 1 \\ & x, y \geq 0. \end{aligned}$$

1. Write its Lagrangian with α, β the multipliers of the inequality constraints.
2. Write the KKT conditions.
3. Use them to solve the problem. For this consider separately the $(\alpha = \beta = 0)$, $(\alpha > 0, \beta = 0)$, $(\alpha = 0, \beta > 0)$, $(\alpha > 0, \beta > 0)$ cases.

Exercise 7
We have worked out the dual problem for the soft SVC problem. Do the same for the simpler **hard** SVC problem

$$\min_{w,b} \frac{1}{2}\|w\|^2$$

subject to $y^p\left(w \cdot x^p + b\right) \geq 1$. What are here the KKT conditions?

Firstly, consider the Lagrangian for this problem

$$\begin{aligned} L(w,b,\lambda) &= \frac{1}{2}\|w\|^2 - \sum \alpha_p \left[y^p\left(w \cdot x^p + b\right) - 1\right] \\ &= \frac{1}{2} w \cdot w - w \sum \lambda_p y^p x^p - b \sum \lambda_p y_p + \sum \lambda_p \\ &= w\left(\frac{1}{2}w - \sum \lambda_p y^p x^p\right) b \sum \lambda_p y_p + \sum \lambda_p \end{aligned}$$

Then, we have to compute the gradient of the Lagrangian

$$\nabla_w L(w,b,\lambda) = w - \sum \lambda_p y^p x^p = 0 \implies w = \sum \lambda_p y^p x^p$$
$$\frac{\partial L}{\partial b} = -\sum \lambda_p y^p = 0$$

Lastly, we have to use the optimalas that we have found in the expression of the Lagrangian:

$$\begin{aligned} L(w,b,\alpha) &= \sum \alpha_p - \frac{1}{2}\left(\sum_p \alpha_p y^p x^p\right)\left(\sum_q \alpha_q y^q x^q\right) \\ &= \sum \alpha_p - \frac{1}{2}\sum_{p,q} \alpha_p y^p \alpha_q y^q x^p x^q = -\alpha \end{aligned}$$

Our optimization problem has turned today

$$\max_{\alpha} \sum \alpha^p - \frac{1}{2} \sum \alpha_p \alpha_q y^p y^q x^p x^q$$

subject to $\alpha^p, \sum \alpha_p y^p = 0$.

Let us state the KKT conditions for this problem.

Exercise 8
A typical Linear Programming (LP) problem can be stated as the following constrained optimization problem:

$$\min_x c \cdot x \quad s.t. \quad x \geq 0, Ax \leq b$$

with $x \in \mathbb{R}^d$, A an $m \times d$ matrix and $b \in \mathbb{R}^m$. A tool often used in LP is to study the so called dual problem, which in this case is

$$\min_z b \cdot z \quad s.t. \quad z \geq 0, A^t z \leq -c$$

with now $z \in \mathbb{R}^m$. Apply our Lagrangian dual construction technique to show that this is indeed the dual formulation of the initial LP problem

Firstly, we have to write the Lagrangian for this problem:

$$L(x, \lambda, \mu) = c \cdot x - \sum_{i=1}^{d} \lambda_i x_i + \sum_{j=1}^{m} \mu_j (a_j \cdot x - b_j)$$

Hence, the gradient respect to x of the lagrangian is:

$$\nabla_x L = c - \lambda + A^t \mu$$

CHECK!!!
(this Lagrangian must have dimension d).

Exercise 11
If Q is a symmetric, positive definite $d \times d$ matrix, show that $f(x) = x^T Q x, \ x \in \mathbb{R}^d$, is a convex function.

Exercise 12
Let $f : \mathbb{R}^d \to \mathbb{R}$ be a function and assume that $epi(f) \subset \mathbb{R}^d \times \mathbb{R}$ is convex. Prove that then f is convex.

Exercise 13

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a convex function. Prove that $\text{epi}(f)$ is a closed set and that $(x, f(x)) \in \partial\, \text{epi}(f)$.

Exercise 14
Prove that if f is strictly convex, it has a unique global minimum.

Exercise 15
Let $f, g : S \subset \mathbb{R}^d \to \mathbb{R}$ be two convex functions on the convex set S . Prove that, as subsets, $\partial(f+g)(x) \subset \partial f(x) + \partial g(x)$ for any $x \in S$.

Exercise 16
Compute the proximal of $f(x) = 0$ and of $g(x) = \frac{1}{2}\|x\|^2$.

Exercise 17
Assume that f is convex. Prove that for any $\lambda > 0, \partial(\lambda f)(x) = \lambda \partial f(x)$ as subsets.

Exercise 19
Compute the proximals of the hinge $f(x) = max\{0, -x\}$ and the ϵ-insensitive $g(x) = max\{0, |x| - \epsilon\}$ loss functions.