Advanced Topics in Statistical Machine Learning: Assignment 2

Mateus Zitelli Dantas

28/06/2016

1 Question 2.12

1.1 a) $\{x \in \mathbb{R} \mid \alpha \leq a^T x \leq b\}$

It is convex because it can be written as the intersection of two half-spaces:

$$\{x \in \mathbb{R} \mid a^T x \leq b\} \bigcap \{x \in \mathbb{R} \mid -a^T x \leq -\alpha\}$$

1.2 b) $\{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq b_i, i = 1, \ldots, n\}$

It is convex because it can be expressed as the intersection of two polyhedrons:

$$\{x \in \mathbb{R}^n \mid x_i \leq -\alpha_i, i = 1, \ldots, n\} \bigcap \{x \in \mathbb{R}^n \mid x_i \leq b_i, i = 1, \ldots, n\}$$

1.3 c) $\{x \in \mathbb{R}^n \mid a_1^T x \leq b_1, a_2^T x \leq b_2\}$

It is convex because it is a polyhedron:

$$\{x \in \mathbb{R}^n \mid Ax \preceq b, A = \begin{bmatrix} a_1^T \\ a_2^T \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}\}$$

1.4 d) $\{x \mid \parallel x - x_0 \parallel_2 \leq \parallel x - y \parallel_2 \forall y \in S\}$

Such set can be expressed as:

$$\bigcap_{y \in S} \{x \mid \parallel x - x_0 \parallel_2 \leq \parallel x - y \parallel_2\}$$

Which is nothing more than intersection of half-spaces. Therefore, it is convex.

1.5 e) $\{x \mid dist(x, S) \leq dist(x, T)\} = C$

It is not convex as showed by the counter example:

$$S = \{0, 2\}, T = \{1\}, x = S_0\theta + (1 - \theta)S_1, \theta = 0.5 \implies x = 1 \notin C.$$

1.6 f) $\{x \mid x + S_2 \subseteq S_1\} = C$

It is convex because it can be expressed as an intersection of convex sets.

$$C = \bigcap_{y \in S_2} \{x \mid x + y \subseteq S_1\} = \bigcap_{y \in S_2} \{x - y \mid x \subseteq S_1\} = \bigcap_{y \in S_2} \{S_1 - y\}$$

1.7 g) $\{x \mid \parallel x - a \parallel_2 \leq \ \theta \parallel x - b \parallel_2\} = C$

$$C = \{x \mid (x - a)^2 \ \leq \ \theta^2 (x - b)^2\}$$

With $\theta = 1$ it became a half-space:

$$C = \{x \mid x^T \leq \frac{(b^T b - a^T a)(b - a)^{-1}}{2}\}$$

With $\theta < 1$, C became an ball:

$$C = \{x \mid (x - x_0)^T (x - x_0) \leq R\}$$

With center and radius:

$$x_0 = \frac{a - \theta^2 b}{1 - \theta^2}, R = \frac{\theta^2 \parallel b \parallel_2^2 - \parallel a \parallel_2^2}{1 - \theta^2} + \parallel x_0 \parallel_2^2$$

2 Question 2.15

2.1 a) $\alpha \leq Ef(x) \leq \beta$

The inequality implies:

$$\alpha \leq \sum_{i=1}^{n} P_i f(a_i) \leq \beta$$

Therefore, P can be interpreted a polyhedron with two linear inequalities as constraints, which makes P convex.

2.2 b) $prob(x > \alpha) \leq \beta$

The inequality implies:

$$\sum_{i: a_i \leq \alpha} P_i \leq \beta$$

It is an affine constraint which makes P convex.

2.3 e) $Ex^2 \geq \alpha$

The inequality implies:

$$\sum_{i=1}^{n} P_i a_i^2 \leq \alpha$$

It is an affine constraint which makes P convex.

3 Question 3.14

3.1 a)

The eigenvalues of the hessian should have opposite signals.

3.2 b)

$$\sup_z \inf_x f(x,z) = \inf_x \sup_z f(x,y)$$

$$\sup_z \inf_x f(x,z) = \sup_z f(\widetilde{x},z) = f(\widetilde{x},\widetilde{z})$$

$$\inf_x \sup_z f(x,z) = \inf_x f(x,\widetilde{z}) = f(\widetilde{x},\widetilde{z})$$

Therefore:

$$\sup_z \inf_x f(x,z) = \inf_x \sup_z f(x,z) = f(\widetilde{x},\widetilde{z})$$

3.3 c)

If $\nabla f(\widetilde{x},\widetilde{z}) \neq 0 \implies \nabla_x f(\widetilde{x},\widetilde{z}) \neq 0 \ or \ \nabla_z f(\widetilde{x},\widetilde{z}) \neq 0$
Suppose a $x' = \widetilde{x} + \nabla_x f(\widetilde{x},\widetilde{z})h$ and $z' = \widetilde{z} + \nabla_z f(\widetilde{x},\widetilde{z})h$ therefore:

$$\exists h \in \mathbb{R} \mid f(\widetilde{x},z') > f(\widetilde{x},\widetilde{z}) \ or \ f(x',\widetilde{z}) < f(\widetilde{x},\widetilde{z})$$

What is an absurd.

4 Question 3.15

4.1 a)

$$\lim_{\alpha\to 0} \frac{x^\alpha - 1}{\alpha}$$

By L'Hospital's Rule it's the same as:

$$\lim_{\alpha\to 0} \frac{\frac{d}{d\alpha}x^\alpha - 1}{\frac{d}{d\alpha}\alpha} = \lim_{\alpha\to 0}\ e^{ln\ x^\alpha} = \lim_{\alpha\to 0}\ x^\alpha ln\ x = ln\ x$$

4.2 b)

if U_α is concave $\nabla^2 U_\alpha \preceq 0$:

$$\nabla^2 U_\alpha = (\alpha - 1)x^{\alpha-2} \leq 0 \ \forall \ 0 \leq \alpha \leq, x \in \mathbb{R}_{++}$$

5 Question 3.16

5.1 a) $f(x) = e^x - 1$

$$\nabla^2 f(x) = e^x \geq 0$$

$$-f(y) \leq -f(x) \implies -e^x(y-x) \leq 0 \implies \nabla - f(x)^T(y-x) \leq 0$$

So the function is convex and quasi-linear.

5.2 b) $f(x_1, x_2) = x_1 x_2$

Seeing that the eigenvalues of $\nabla^2 f(x)$ are 1 and -1, so the function is neither convex or concave.
Verifying by the second-order condition for quasi-convexity:

$$y^* \nabla f = 0 \implies y_1^* x_2^* = -y_2^* x_1^*$$

$$y^{*T} \nabla^2 f y^* = 2y_1^* y_2^* = 2y_1^* y_2^* = -2x_1^* y_2^{*2}$$

Seeing that $y^{*T} \nabla^2 f y^* \leq 0$ in $\mathbb{R}^2_{++}$ this function is quasi-concave.

5.3 c) $f(x) = \frac{1}{x_1 x_2}$

$$|\nabla^2 f| = 3 * x_1^{-4} x_2^{-4} \geq 0$$

So the function is convex and quasi-convex.

5.4 d) $f(x_1, x_2) = \frac{x_1}{x_2}$

$$|\nabla^2 f| = -x_2^{-4} < 0$$

So it is not convex neither concave and both:

$$\{x \in dom f \mid \frac{x_1}{x_2} < \alpha\}$$

$$\{x \in dom f \mid \frac{x_1}{x_2} > \alpha\}$$

Are half-spaces, so the function is quasi-linear.

5.5 d) $f(x_1, x_2) = \frac{x_1^2}{x_2^{-1}}$

By the Silvester's Criterion:

$$2x_2^{-1} \geq 0$$

$$4x_1^2 x_2^{-4} - 4x_1^2 x_2^{-4} \geq 0$$

Therefore:

$$0 \preceq \nabla^2 f(x)$$

So it's convex and quasi-convex.

5.6 d) $f(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha}$

$$\nabla^2 f(x) = -\alpha(1-\alpha) x_1^{\alpha} x_2^{1-\alpha} \begin{bmatrix} x_1^{-1} \\ -x_2^{-1} \end{bmatrix} \begin{bmatrix} x_1^{-1} & -x_2^{-1} \end{bmatrix} \preceq 0$$

It is concave and quasi-concave.

6 Question 3.36

6.1 a) $f(x) = \max_i x_i$

Supposing $x_1 = \cdots = x_n = t$

$$\sup_x (y^T x - \max_i x_i) \geq t(\sum_i y_i - 1)$$

if $\sum_i y_i \neq 1$ implies $f^*(y)$ unbounded.
Case $1^T y = 1$, we have:

$$f^*(y) = \sup_x (\sum_i y_i (x_i - x_{[1]}))$$

if $y_i \geq 0 \implies f^*(y) = 0$, otherwise $f^*(y) = \infty$, therefore:

$$f^*(y) = \begin{cases} 0 & 1^T y = 1 \text{ and } y \succeq 0 \\ \infty & \text{otherwise} \end{cases}$$

6.2 b) $f(x) = \sum_{i=1}^{r} x_{[i]}$

Suppose $x_1 = \cdots = x_n = t$ then:

$$\sup_x (y^T x - \sum_{i=1}^{r} x_{[i]}) \geq t(\sum_{i=1}^{n} y_i - r)$$

if $\sum_i y_i \neq r$ implies $f^*(y)$ unbounded.
Case $1^T y = r$, we have:

$$f^*(y) = \sup_x(\sum_i^n y_i x_i - \sum_i^r x_{[i]}) = \sup_x(\sum_i^r (y_{[i]}-1)x_i + \sum_{j=r+1}^n x_{[j]}y_{[j]})$$

Supposing $0 \preceq y \preceq 1$ and $x_{[i]} \geq x_{[r]} \geq x_{[j]}$, therefore $x_{[j]}y_{[j]} \leq x_{[r]}y_{[j]}$ and $(y_{[i]}-1)x_{[i]} \leq (y_{[i]}-1)x_{[r]}$. With that:

$$f^*(y) \leq \sup_x(\sum_{i=1}^r (y_{[i]}-1)x_{[r]} + x_{[r]} \sum_{j=r+1}^n y_{[j]}) = 0$$

Therefore:

$$f^*(y) = \begin{cases} 0 & 1^T y = r \text{ and } 0 \preceq y \preceq 1 \\ \infty & \text{otherwise} \end{cases}$$

6.3 e) $f(x) = -(\prod_{i=1}^n x_i)^{1/n}$

If $y_i > 0$ then setting $x_i \to \infty$ $f^*(y)$ is unbounded. In case of $y \preceq 0$ we can use the arithmentic-geometric means inequality to find the domain of $f^*(y)$. Suppose $k = -y \succeq 0$.

$$-y^T x = k^T x = n(\frac{1}{n}\sum_i^n k_i x_i) \geq n(\prod_i^n k_i x_i)^{\frac{1}{n}} = n(\prod_i^n k_i)^{\frac{1}{n}}(\prod_i^n x_i)^{\frac{1}{n}} = n(\prod_i^n k_i)^{\frac{1}{n}} f(x)$$

Therefore:

$$y^T x \leq n(\prod_i^n y_i)^{\frac{1}{n}} f(x)$$

With that we can say:

$$f^*(y) \leq n(\prod_i^n y_i)^{\frac{1}{n}} f(x) - f(x) = f(x)(n(\prod_i^n y_i)^{\frac{1}{n}} - 1)$$

Therefore, to constrain $f^*(y)$:

$$(\prod_i^n y_i)^{\frac{1}{n}} \leq \frac{1}{n}$$

$$f^*(y) = \begin{cases} 0 & y \preceq 0, (\prod_i^n y_i)^{\frac{1}{n}} \leq \frac{1}{n} \\ \infty & \text{otherwise} \end{cases}$$

7 Polyhedron

$$A = \begin{bmatrix} -1 & 1 \\ 3 & 2 \\ 1 & 0 \\ -2 & -3 \\ -1 & 0 \end{bmatrix}, b = \begin{bmatrix} 3 \\ 6 \\ 2 \\ 2 \\ 1 \end{bmatrix}$$

Where the polyhedron is $\{x \mid Ax \preceq b\}$