Tarea I - Optimización no lineal

Solución 1. *Notemos que la función $f : X \to \overline{\mathbb{R}}$ esta bien definida, pues en la recta real extendida siempre existe el supremo. Recordemos que dado $\gamma \in \mathbb{R}$ se define el sub nivel asociado como*

$$\Gamma_\gamma(f) := f^{-1}((-\infty, \gamma]) = \{x \in X \mid f(x) \leq \gamma\}$$

Ahora notemos lo siguiente

$$f(x) \leq \gamma \iff \sup_{\alpha \in \mathcal{A}} f_\alpha(x) \leq \gamma \iff \forall \alpha \in \mathcal{A}, f_\alpha(x) \leq \gamma \tag{1}$$

De lo que concluimos que

$$\Gamma_\gamma(f) = \bigcap_{\alpha \in \mathcal{A}} \Gamma_\gamma(f_\alpha)$$

Dado que cada f_α es τ-s.c.i. tenemos que $\Gamma_\gamma(f_\alpha)$ es cerrado para todo $\alpha \in \mathcal{A}$. Dado que la intersección arbitraria de cerrados es cerrada, tenemos que $\Gamma_\gamma(f)$ es cerrado para todo $\gamma \in \mathbb{R}$, es decir f es τ-s.c.i.

Ahora supongamos que las f_α son convexas para todo $\alpha \in \mathcal{A}$. Por lo tanto tenemos que el epigrafo es convexo para toda f_α. Recordemos que el epigrafo se define como

$$\text{epi}(f) := \{(x, \gamma) \in X \times \mathbb{R} \mid f(x) \leq \gamma\}$$

Por la ecuación (1), *tenemos la siguiente igualdad*

$$\text{epi}(f) = \bigcap_{\alpha \in \mathcal{A}} \text{epi}(f_\alpha)$$

Dado que la intersección (arbitraria) de convexos es convexa, tenemos que epi(f) *es convexo y por tanto f es convexa.*

Solución 2. *Directo, pues las funciones continuas son semicontinuas inferior, esto es facil de ver pues tenemos la igualdad*

$$\Gamma_\gamma(f) = f^{-1}((-\infty, \gamma])$$

Dado que preimagen de cerrados bajo funciones continuas es cerrado, tenemos que si f es continua, el subnivel es cerrado.

Luego la funcion $x \mapsto ||Ax - b||_Y$ es continua, pues $x \mapsto Ax - b$ es continua para $b \in Y$ (pues A es continua). Dado que $|| \cdot ||_Y : Y \to \mathbb{R}$ es continua, la composición tambien lo es.

Solución 3. *Notemos que*

$$\forall x \in X, \text{Val}(P) \leq f(x) \implies \mu(S)\,\text{Val}(P) \leq \int_S f d\mu$$

Dado que $\mu(S) = 1$, obtenemos que para toda medida $\mu \geq 0$ tal que $\mu(S) = 1$ se tenga la desigualdad. De esto obtenemos que

$$\text{Val}(P) \leq \text{Val}(P_m)$$

Si consideramos una sucesion minimizante $\{x_n\} \subseteq X$ *tal que* $f(x_n) \to \mathrm{Val}(P)$ *nos construimos la sucesion de medidas* $\{\delta_{x_n}\} \subseteq \mathcal{M}^+$, *donde* $\delta_{x_n}(A) = 1 \iff x_n \in A$, *obtenemos que*

$$\int_S f d\delta_{x_n} = f(x_n) \underset{n\to\infty}{\to} \mathrm{Val}(P)$$

Por lo tanto $\mathrm{Val}(P) = \mathrm{Val}(P_m)$

Solución 4.

1. *En efecto, sea* $A, B \in \mathbf{S}^n$, *luego tenemos que*

$$\mathrm{Tr}(AB) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ji}$$

Comprobemos que en efecto es un producto interno

$$\mathrm{Tr}(A^2) = 0 \iff \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}a_{ji} = 0 \iff \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2 = 0 \iff \forall i,j \in \{1,\dots,n\} a_{ij} = 0 \iff A = 0$$

Ademas notemos que de lo anterior tenemos que

$$\langle A, A\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2 \geq 0$$

Veamos que es simetrico

$$\langle A, B\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ji} = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}a_{ji} = \langle B, A\rangle$$

Verifiquemos linealidad en el primer argumento, pues la bilineallidad sigue de esto y la simetria. Sea $C \in \mathbf{S}$ *y* $\lambda \in \mathbb{R}$, *luego*

$$\langle A + \lambda B, C\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n}(a_{ij} + \lambda b_{ij})c_{ji} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}c_{ji} + \lambda \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}c_{ji} = \langle A, C\rangle + \lambda\langle B, C\rangle$$

Por lo tanto la traza si define un producto interno. Dado que el espacio es de dimension finita tenemos que es completo y por tanto un espacio de Hilbert.

2. *Notemos que de la definicion de* P, *la matriz es claramente simetrica pues*

$$P_{ij} = x_i^T x_j = x_j^T x_i = P_{ji}$$

Veamos que $P = X^T X$. *Notemos que*

$$(X^T X)_{ij} = \sum_{k=1}^{n} x_{k,i}x_{k,j} = x_i^T x_j$$

Por lo tanto se tiene la igualdad $P = X^T X$. *Si* $b \in \mathbb{R}^n$, *tenemos que*

$$b^T P b = b^T X^T X b = (Xb)^T X b = ||Xb||^2 \geq 0$$

3. *Notemos primero la siguiente igualdad*

$$\sum_{i,j=1}^{n} C_{ij}\left(\frac{1 - x_i^T x_j}{2}\right) = \frac{1}{2}\sum_{i,j=1}^{n} C_{ij} - \frac{1}{2}\operatorname{Tr}(CX^TX)$$

Luego maximizar la suma es lo mismo que minimizar $\operatorname{Tr}(CX^TX)$. *Luego esto es equivalente a minimizar* $\operatorname{Tr}(CM)$ *con* $M \in S_+^n(\mathbb{R})$, *pues toda matriz simetrica definida semi-positiva tiene una descompsocion de Cholesky, es decir* $M = X^TX$ *para alguna matriz* X. *Luego dado que* M *es de Gram, sabemos que en la diagonal estan las normas de los elementos de las columnas de su descomposicion. Por lo tanto lo que queremos es*

$$\textit{Minimizar } \operatorname{Tr}(CM) \textit{ con } M \in S_+^n(\mathbb{R})$$
$$\textit{tal que } \operatorname{Tr}(A_iM) = 1 \textit{ donde } A_i \textit{ tiene el vector canonico } e_i \textit{ en la columna } i \textit{ para } i \in \{1, \dots, n\}$$

Solución 5. $(\Leftarrow)$ *Supongamos que para todo* $\{x_1, \dots, x_n\} \subseteq X$ *y* $\lambda_1, \dots, \lambda_n \in [0,1]$ *se tiene que*

$$\sum_{i=1}^{n} \lambda_i = 1 \implies f\Big(\sum_{i=1}^{n} \lambda_i x_i\Big) \leq \sum_{i=1}^{n} f(x_i)$$

Luego dado $x, y \in X$ *y* $\lambda \in [0,1]$, *tenemos que* $\lambda + (1-\lambda) = 1$, *por lo tanto por hipotesis tenemos que*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$

Es decir, f *es convexa.*

$(\implies)$ *Procedamos por inducción. El caso para* $n = 2$ *es trivial, pues es la convexidad usual. Supongamos para* n *y demostramos para* $n+1$, *Sean* $x_1, \dots, x_{n+1} \in X$ *y* $\lambda_1, \dots, \lambda_{n+1} \in [0,1]$ *tal que sumen* 1. *Luego tenemos que*

$$f\Big(\lambda_{n+1}x_{n+1} + (1-\lambda_{n+1})\Big(\sum_{k=1}^{n} \lambda_k x_k\Big)\Big) \leq \lambda_{n+1} f(x_{n+1}) + (1-\lambda_{n+1}) f\Big(\sum_{k=1}^{n} \frac{\lambda_k}{1-\lambda_{n+1}} x_k\Big)$$

Dado que los coeficientes de la derecha suman 1, *pues* $1 - \lambda_{n+1} = \sum_{k=1}^{n} \lambda_k$, *aplicamos la hipotesis de induccion y obtenemos que*

$$f\Big(\sum_{k=1}^{n+1} \lambda_k x_k\Big) \leq \lambda_{n+1} f(x_{n+1}) + \sum_{k=1}^{n} \lambda_k f(x_k) = \sum_{k=1}^{n+1} \lambda_k f(x_k)$$