Tarea Tres
Optimización I

Esteban Reyes Saldaña

15 de febrero de 2021

Problema 1. ¿El conjunto $S = \{a \in \mathbb{R}^K | p(0) = 1, |p(t)| \leq 1 \text{ para } t \in [\alpha, \beta]\}$ donde

$$p(t) = a_1 + a_2 t + \cdots + a_k t^{k-1},$$

es convexo?

Solución. Sí es convexo. Sean $a, b \in S$ entonces $a, b \in \mathbb{R}^k$. Por cerradura,

$$\alpha a + (1-\alpha) b \in \mathbb{R}^K$$

para $\alpha \in \mathbb{R}^k$. Luego,

$$\begin{aligned} \alpha a + (1-\alpha) b &= \alpha [a_1, a_2, \ldots, a_k]^T - (1-\alpha)[b_1, b_2, \ldots, b_k]^T \\ &= [\alpha a_1 + (1-\alpha) b_1, \ldots, \alpha a_k + (1-\alpha) b_k] \end{aligned}$$

Como $a \in S$,

$$p(0) = a_1 = 1.$$

Como $b \in S$,

$$p(0) = b_1 = 1.$$

Ahora,

$$\begin{aligned} p(0) &= \alpha a_1 + (1-\alpha) b_1 \\ &= \alpha + (1-\alpha) \\ &= 1. \end{aligned}$$

Por otro lado,

$$\begin{aligned} 1 &\geq |p_a(t)| \\ &= |a_1 + a_2 t + \cdots + a_k t^{k-1}| \end{aligned}$$

y

$$\begin{aligned} 1 &\geq |p_b(t)| \\ &= |b_1 + b_2 t + \cdots + b_k t^{k-1}| \end{aligned}$$

Así que

$$\begin{aligned} |p_{\alpha a+(1-\alpha)b}(t)| &= |\alpha p_a(t) + (1-\alpha)p_b(t)| \\ &\leq |\alpha||p_a(t)| + |1-\alpha||p_b(t)| \\ &\leq \alpha(1) + (1-\alpha)(1) \\ &= 1 \end{aligned}$$

Así que $p_{\alpha a+(1-\alpha)b}(t) \in S$. Se concluye que S es convexo.

Problema 2. Suponga que f es convexa, $\lambda_1 > 0$ y $\lambda_2 \leq 0$ con $\lambda_1 + \lambda_2 = 1$ y $x_1, x_2 \in dom(f)$. Demuestre que la desigualdad

$$f(\lambda_1 x_1 + \lambda_2 x_2) \geq \lambda_1 f(x_1) + \lambda_2 f(x_2)$$

siempre es verdadera.

Demostración. Como $\lambda_2 \leq 0$ y $-\lambda_2 \geq 0$ entonces

$$\lambda_1 = 1 - \lambda_2 \geq 1$$

así que

$$0 < \frac{1}{\lambda_1} \leq 1. \tag{1}$$

Además,

$$\begin{aligned} -1 &\leq -\frac{1}{\lambda_1} < 0 \\ 0 &\leq 1 - \frac{1}{\lambda_1} < 1. \end{aligned}$$

Notemos que

$$\begin{aligned} \frac{1}{\lambda_1}(\lambda_1 x_1 + \lambda_2 x_2) + \left(1 - \frac{1}{\lambda_1}\right) x_2 &= x_1 + \frac{\lambda_2}{\lambda_1} x_2 + \left(\frac{\lambda_1 - 1}{\lambda_1}\right) x_2 \\ &= x_1 + \frac{\lambda_2}{\lambda_1} x_2 - \frac{\lambda_2}{\lambda_1} x_2 \\ &= x_1. \end{aligned}$$

Además, por hipótesis

$$\lambda_1 x_1 + \lambda_2 x_2 \in Dom(f).$$

Luego, dado que $f(x)$ es convexa y usando $\dfrac{1}{\lambda_1} \in (0,1]$ entonces

$$\begin{aligned}
f(x_1) &= f\left(\frac{1}{\lambda_1}(\lambda_1 x_1 + \lambda_2 x_2) + \left(1 - \frac{1}{\lambda_1}\right) x_2\right) \\
&\leq \frac{1}{\lambda_1} f(\lambda_1 x_1 + \lambda_2 x_2) + \left(1 - \frac{1}{\lambda_1}\right) f(x_2) \\
&= \frac{1}{\lambda_1} f(\lambda_1 x_1 + \lambda_2 x_2) - \frac{\lambda_2}{\lambda_1} f(x_2) \\
\lambda_1 f(x_1) &\leq f(\lambda_1 x_1 + \lambda_2 x_2) - \lambda_2 f(x_2) \\
f(\lambda_1 x_1 + \lambda_2 x_2) &\geq \lambda_1 f(x_1) + \lambda_2 f(x_2).
\end{aligned}$$

□

Problema 3. Demuestre que la función $f : \mathbb{R}^n \to \mathbb{R}$

$$f(x) = \exp(-g(x))$$

es convexo. Donde $g : \mathbb{R}^n \to \mathbb{R}$ tiene dominio convexo y satisface que

$$\begin{bmatrix} \nabla^2 g(x) & \nabla g(x) \\ \nabla^T g(x) & 1 \end{bmatrix} \geq 0$$

para $x \in dom(g)$.

Demostración. En clase se vió que una función es convexa si y sólo si su Hessiano es definido semipositivo sobre su dominio covexo. Entonces se probará que $-exp(-g(x))$ es convexa usando este resultado.
Utilizando la relación gradiente-derivada y regla de la cadena tenemos que

$$\begin{aligned}
\nabla f(x) &= -exp(-g(x))(-1)\nabla g(x) \\
&= exp(-g(x))\nabla g(x).
\end{aligned}$$

Para calcular el Hessiano se usa lo anterior y la regla de producto, entonces

$$\begin{aligned}
\nabla^2 f(x) &= \nabla(exp(-g(x))\nabla g(x)) \\
&= exp(-g(x))\nabla^2 g(x) + \nabla g(x) \exp(-g(x))(-1)\nabla^T g(x) \\
&= exp(-g(x)) \left[\nabla^2 g(x) - \nabla g(x)\nabla^T g(x)\right]
\end{aligned}$$

Sabemos que Por otro lado, dado que la matriz del enunciado del problema está definida semi positiva entonces para todo $x \in \mathbb{R}^n$

$$xTAx \geq 0$$

en particular para $u^T = [x^T, -x^T \nabla g(x)]$. Luego

$$\begin{aligned}
0 &\leq u^T \begin{bmatrix} \nabla^2 g(x) & \nabla g(x) \\ \nabla^T g(x) & 1 \end{bmatrix} u \\
&= [x^T, -x^T \nabla g(x)] \begin{bmatrix} \nabla^2 g(x) & \nabla g(x) \\ \nabla^T g(x) & 1 \end{bmatrix} \begin{bmatrix} x \\ -\nabla^T g(x) x \end{bmatrix} \\
&= [x^T, -x^T \nabla g(x)] \begin{bmatrix} \nabla^2 g(x) x - \nabla g(x) \nabla^T g(x) x \\ \nabla^T g(x) x - \nabla^T g(x) x \end{bmatrix} \\
&= x^T \nabla^2 g(x) x - x^T \nabla g(x) \nabla^T g(x) x \\
&\quad -x^T \nabla^T g(x) x + x^T \nabla g(x) \nabla^T g(x) x \\
&= x^T \nabla^2 g(x) x - x^T \nabla^T g(x) x.
\end{aligned}$$

y sabemos que $exp(-g(x))$ es mayor que cero. Entonces

$$\nabla^2 f(x) = exp(-g(x)) \left[\nabla^2 g(x) - \nabla g(x) \nabla^T g(x)\right] \geq 0.$$

De lo anterior, $\nabla^2 f(x) \geq 0$ y entonces es definida semipositiva. Luego, concluímos que $f(x)$ es convexa. □

Problema 4. Demuestre que $f(x, y) = x^2/y$, $y > 0$ es convexo.

Demostración. De nuevo se probará que $f(x, y)$ es convexa probando que su Hessiano es siempre definido positivo. Por un lado,

$$\begin{aligned}
\nabla f(x, y) &= \left[2x/y, x^2(-1)/y^2\right] \\
&= \left[2x/y, -x^2/y^2\right].
\end{aligned}$$

Entonces,

$$\begin{aligned}
\nabla^2 f(x, y) &= \begin{bmatrix} 2/y & 2x(-1)/y^2 \\ -2x/y^2 & -x^2(-2)/y^3 \end{bmatrix} \\
&= \begin{bmatrix} 2/y & -2x/y^2 \\ -2x/y^2 & 2x^2/y^3 \end{bmatrix} \\
&= \frac{2}{y^3} \begin{bmatrix} y^2 & -xy \\ -xy & x^2 \end{bmatrix}
\end{aligned}$$

Como $y > 0$ entonces
$$\frac{2}{y^3} > 0. \tag{2}$$
Por otro lado, sea $s = [s_1, s_2]^T \in \mathbb{R}^2$. Luego,

$$\begin{aligned}
s^T \nabla^2 f(x,y) s &= [s_1, s_2] \frac{2}{y^3} \begin{bmatrix} y^2 & -xy \\ -xy & x^2 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \\
&= \frac{2}{y^3} [s_1, s_2] \begin{bmatrix} y^2 s_1 - xy s_2 \\ -xy s_1 + x^2 s_2 \end{bmatrix} \\
&= \frac{2}{y^3} [(y^2 s_1 - xy s_2) s_1 + (x^2 s_2 - xy s_1) s_2] \\
&= \frac{2}{y^3} [y^2 s_1^2 - xy s_1 s_2 + x^2 s_2^2 - xy s_1 s_2] \\
&= \frac{2}{y^3} [y^2 s_1^2 - 2xy s_1 s_2 + x^2 s_2^2] \\
&= \frac{2}{y^3} (y s_1 - x s_2)^2
\end{aligned}$$

Ahora, dado (2) y el hecho que , en $\mathbb{R}$, todo número al cuadrado es mayor o igual que cero entonces

$$s^T \nabla^2 f(x,y) s = \frac{2}{y^3} (y s_1 - x s_2)^2 \geq 0.$$

Por lo tanto $\nabla^2 f(x,y)$ es definida semipositiva. De lo anterior se concluye que $f(x,y)$ es convexa. □

Problema 5. Considere la función $f(x_1, x_2) = (x_1 + x_2^2)^2$. En el punto $x^T = [1, 0]$ considere la dirección de búsqueda $p^T = [-1, 1]$. Demuestre que p es una dirección de descenso y encuentre todos los minimizadores de la función.

Demostración. Recordemos que p es una dirección de descenso si

$$\nabla^T f(x) p < 0.$$

Por un lado,

$$\begin{aligned}
\nabla f(x_1, x_2) &= \left[2(x_1 + x_2^2), 2(x_1 + x_2^2)(2x_2)\right]^T \\
&= \left[2(x_1 + x_2^2), 4(x_1 + x_2^2) x_2\right]^T
\end{aligned}$$

entonces

$$\begin{aligned}\nabla f(1,0) &= [2(1+0), 2(1+0)0]^T \\ &= [2,0]^T\end{aligned}$$

Así que

$$\begin{aligned}\nabla f(1,0)p &= [2,0]^T \begin{bmatrix}-1 \\ 1\end{bmatrix} \\ &= -2+0 \\ &< 0.\end{aligned}$$

De donde se concluye que p es una dirección de descenso para $f(x_1, x_2)$.

(i) *Minimizadores usando gradiente-Hessiano*. El Hessiano de $f(x_1, x_2)$ es

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix}2x_2^2 & 4x_2 \\ 4x_2 & 4(x_1 + 3x_2^2)\end{bmatrix} \tag{3}$$

Los puntos tales que

$$\nabla f(x_1, x_2) = 0$$

deben cumplir

$$\begin{cases} x_1 + x_2^2 = 0 \\ (x_1 + x_2^2)x_2 = 0. \end{cases} \tag{4}$$

entonces $x_1 = x_2 = 0$. Luego,

$$\nabla^2 f(0,0) = \begin{bmatrix}0 & 0 \\ 0 & 0\end{bmatrix} \tag{5}$$

por lo que no podemos decir si dicho punto es mínimo.

(ii) *Minimizadores en dirección x y p.*

Sea $g(\alpha) : \mathbb{R} \to \mathbb{R}$ dada por

$$g(\alpha) = f(x + \alpha p).$$

Queremos encontrar todos los minimizadores de $g(\alpha)$ para $\alpha > 0$. Entonces buscamos aquellos puntos en los que

$$\begin{aligned}0 &= g'(\alpha) \\ &= \nabla^T f(x + \alpha p)p.\end{aligned}$$

con $x^T = [1, 0]$ y $p^T = [-1, 1]$. Así que, para esta configuración tenemos

$$\begin{aligned} x + \alpha p &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -\alpha \\ \alpha \end{bmatrix} \\ &= \begin{bmatrix} 1 - \alpha \\ \alpha \end{bmatrix}. \end{aligned}$$

Ahora,

$$\begin{aligned} 0 &= \nabla^T f(x + \alpha p) \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ &= [2(1 - \alpha + \alpha^2), 4\alpha(1 - \alpha + \alpha^2)] \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ &= -2(1 - \alpha + \alpha^2) + 4\alpha(1 - \alpha + \alpha^2) \\ &= (4\alpha - 2)(1 - \alpha + \alpha^2) \end{aligned}$$

Notemos que

$$\begin{aligned} \alpha^2 - \alpha + 1 &= (\alpha^2 - \alpha + 1/4 + 3/4 \\ &= (\alpha - 1/2)^2 + 3/4 \\ &> 0. \end{aligned}$$

Así que $(4\alpha - 2)(1 - \alpha + \alpha^2) = 0$ si y sólo si

$$\alpha = \frac{1}{2}.$$

Dado que ya se desmostró que p es una dirección de descenso y que $f(x+\alpha p)$ crece cuando α crece, se concluye que

$$x + \alpha p = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$$

es el único minimizador de $f(x_1, x_2)$. $\square$

Problema 6. Encuentre todos los valores del parámetro a tales que $[1, 0]^T$ es el minimizador o maximizador de la función

$$f(x_1, x_2) = a^3 x_1 e^{x_2} + 2a^2 \log(x_1 + x_2) - (a + 2)x_1 + 8ax_2 + 16x_1x_2.$$

Solución. Dado que buscamos los puntos tales que $[1,0]^T$ es minimizador o maximizador entonces se debe cumplir que

$$\nabla f(1,0) = 0.$$

Por un lado,

$$\nabla f(x_1,x_2) = \begin{bmatrix} a^3e^{x_2} + \dfrac{2a^2}{x_1+x_2} - (a+2) + 16x_2 \\ a^3x_1e^{x^2} + \dfrac{2a^2}{x_1+x_2} + 8a + 16x_1 \end{bmatrix}$$

Luego,

$$\begin{aligned} \nabla f(1,0) &= \begin{bmatrix} a^3 + 2a^2 - (a+2) \\ a^3 + 2a^2 + 8a + 16 \end{bmatrix} \\ &= \begin{bmatrix} a^3 + 2a^2 - a - 2 \\ a^3 + 2a^2 + 8a + 16 \end{bmatrix} \end{aligned}$$

Así que el gradiente de $f(x_1,x_2)$ es nulo si se cumplen al mismo tiempo

$$\begin{aligned} 0 &= a^3 + 2a^2 - a - 2 = (a+2)(a+1)(a-1) \\ 0 &= a^3 + 2a^2 + 8a + 16 = (a+2)(a^2+8) \end{aligned}$$

por lo que la única solución es

$$a = -2.$$

Se probará ahora que dicho punto no corresponde a un punto silla. El Hessiano de $f(x_1,x_2)$ es

$$\nabla^2 f(x_1,x_2) = \begin{bmatrix} -\dfrac{2a^2}{(x1+x_2)^2} & a^3e^{x_2} - \dfrac{2a^2}{(x_1+x_2)} + 16 \\ a^3e^{x_2} - \dfrac{2a^2}{(x_1+x_2)} + 16 & a^3e^{x_2} - \dfrac{2a^2}{(x_1+x_2)^2} \end{bmatrix}$$

Para $a = -2$ y $x^T = [1, 0]$ se tiene

$$
\begin{aligned}
\nabla^2 f(1,0) &= \begin{bmatrix} -\dfrac{2(-2)^2}{(1+0)^2} & (-2)^3e^0 - \dfrac{2(-2)^2}{(1+0)} + 16 \\ (-2)^3e^0 - \dfrac{2(-2)^2}{(1+0)^2} + 16 & (-2)^3e^0 - \dfrac{2(-2)^2}{(1+0)^2} \end{bmatrix} \\
&= \begin{bmatrix} -8 & -8-8+16 \\ -8-8+16 & -8-8 \end{bmatrix} \\
&= \begin{bmatrix} -8 & 0 \\ 0 & -16 \end{bmatrix}
\end{aligned}
$$

cuyos valores propios son todos negativos y por lo tanto, dicho punto corresponde a un máximo.

Problema 7. Considere la sucesión $x_k = 1 + 1/k!$, $k = 0, 1, \ldots$ ¿Converge linealmente a 1? Justifique su respuesta.

Solución. Sabemos que

$$\lim_{k\to\infty} \frac{1}{k!} = 0.$$

Entonces

$$
\begin{aligned}
\lim_{k\to\infty} x_k &= \lim_{k\to\infty} \left(1 + \frac{1}{k!}\right) \\
&= 1 \\
&= x^*.
\end{aligned}
$$

Ahora,

$$
\begin{aligned}
\frac{||x_{k+1} - x^*||}{||x_k - x^*||} &= \frac{||1 + \dfrac{1}{(k+1)!} - 1||}{||1 + \dfrac{1}{k!} - 1||} \\
&= \frac{\dfrac{1}{(k+1)!}}{\dfrac{1}{k!}} \\
&= \frac{k!}{(k+1)!} \\
&= \frac{1}{k+1} \\
&\to 0, \text{ si } k \to \infty.
\end{aligned}
$$

Por lo que $\{x_k\}$ converge linealmente a 1.

Problema 8. Demuestre que $f(x) = \log\left(\sum_{i=1}^{n} \exp(x_i)\right)$ es convexa.

Demostración. Primero veamos que si

$$f(x) = log(\phi(x))$$

tenemos que

$$\nabla f(x) = \frac{1}{\phi(x)} \nabla \phi(x)$$

Ahora,

$$\begin{aligned} \nabla^2 f(x) &= -\frac{1}{\phi^2(x)} \nabla^T \phi(x) \nabla \phi(x) + \frac{1}{\phi(x)} \nabla^2 \phi(x) \\ &= \frac{1}{\phi^2(x)} (\phi(x) \nabla^2 \phi(x) - \nabla^T \phi(x) \nabla \phi(x)) \end{aligned}$$

Como

$$\phi(x) = \sum_{i=1}^{n} exp(x_i) > 0,$$

para verificar que $\nabla^2 f(x)$ es definido positivo entonces basta verificar que

$$A = (\phi(x) \nabla^2 \phi(x) - \nabla^T \phi(x) \nabla \phi(x))$$

es definida positiva.
Para la función dada

$$\phi(x) = \sum_{i=1}^{n} \exp(x_i)$$

si $z = [\exp(x_1), \ldots, \exp(x_n)]^T$ entonces

$$\nabla \phi(x) = z$$

y

$$\nabla^2 \phi(x) = diag(z).$$

Sea $v \in \mathbb{R}^n$ entonces

$$\begin{aligned} v^T A v &= v^T[\phi(x)\nabla^2\phi(x) - \nabla^T\phi(x)\nabla\phi(x)]v \\ &= v^T[\phi(x) diag(z) - z^T z]v \\ &= v^T\phi(x) diag(z)v - v^T z^T z v \\ &= \phi(x) v^T diag(z) v - v^T z^T z v \\ &= \phi(x)\left(\sum_{i=1}^n \exp(x_i) v_i^2\right) - \left(\sum_{i=1}^n v_i \exp(x_i)\right)^2 \end{aligned}$$

Usando la desigualdad de Cauchy-Swartz tenemos que

$$\left(\sum_{i=1}^n v_i \exp(x_i)\right)^2 \leq \sum_{i=1}^n \exp(x_i)\left(\sum_{i=1}^n \exp(x_i) v_i^2\right)$$

Por lo que $\nabla^2 f(x)$ es definida positiva y por lo tanto, $f(x)$ es convexa. □

Problema 9. Demuestre que $f(x) = \log\left(\sum_{i=1}^n \exp(g_i(x))\right) : \mathbb{R} \to \mathbb{R}$ es convexa si $g_i : \mathbb{R} \to \mathbb{R}$ es convexa.

Demostración. Sea $\alpha \in (0,1]$ y $x, y \in \mathbb{R}$. Sean $x_1, x_2, \ldots, x_n \in \mathbb{R}^+$ entonces

$$\sum_{i=1}^n x_i^\alpha = \left(\sum_{i=1}^n x_i\right)^\alpha \tag{6}$$

Luego,

$$
\begin{aligned}
f(\alpha x + (1-\alpha)y) &= \log\left(\sum_{i=1}^{n} \exp(g_i(\alpha x + (1-\alpha)y))\right) \\
&\leq \log\left(\underbrace{\sum_{i=1}^{n} \exp(\underbrace{\alpha g_i(x) + (1-\alpha)g_i(y)}_{\text{convexidad de } g_i})}_{\text{monotonía de } exp}\right) \\
&= \log\left(\sum_{i=1}^{n} \exp(\alpha g_i(x)) \exp((1-\alpha)g_i(y))\right) \\
&= \log\left(\sum_{i=1}^{n} \exp(g_i(x))^{\alpha} \exp(g_i(y))^{1-\alpha}\right) \\
&\leq \log\left(\left(\sum_{i=1}^{n} \exp(g_i(x))\right)^{\alpha} \left(\sum_{i=1}^{n} \exp(g_i(x))\right)^{1-\alpha}\right) \\
&= \log\left(\left(\sum_{i=1}^{n} \exp(g_i(x))\right)^{\alpha}\right) + \log\left(\left(\sum_{i=1}^{n} \exp(g_i(y))\right)^{1-\alpha}\right) \\
&= \alpha \log\left(\sum_{i=1}^{n} \exp(g_i(x))\right) + (1-\alpha)\log\left(\sum_{i=1}^{n} \exp(g_i(y))\right) \\
&= \alpha f(x) + (1-\alpha) f(y).
\end{aligned}
$$

Por lo tanto $f(x)$ es convexa. □

Problema 10. Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función diferenciable. Demuestre que f es convexa sobre un conjunto convexo no vacío C si y sólo si

$$(\nabla f(x) - \nabla f(y))^T (x - y) \geq 0, \forall x, y \in C$$

Nota: la prueba que tenemos es solo para el caso $(\to)$.

Demostración. Suponga que

$$(\nabla f(x) - \nabla f(y))^T (x - y) \geq 0, \forall x, y \in C$$

en particular, si $\alpha \in (0,1]$ y se toman $x, y \in C$, se tiene que

$$x + \alpha(y - x) \in C$$

y así

$$\begin{aligned} 0 &\leq (\nabla f(x + \alpha(y-x)) - \nabla f(x))^T (x + \alpha(y - x) - x) \\ &= \alpha(\nabla f(x + \alpha(y-x)) - \nabla f(x))^T (y - x) \end{aligned}$$

Dado que, por construcción, $\alpha > 0$ entonces

$$\nabla^T f(x + \alpha(y-x))(y-x) \geq \nabla^T f(x)(y-x) \tag{7}$$

Ahora, sea

$$\phi(\alpha) = f(x + \alpha(y-x)), \alpha \in (0,1].$$

Notemos que

$$\phi'(\alpha) = \nabla^T f(x + \alpha(y-x))(y-x). \tag{8}$$

y además

$$\begin{aligned} \phi(1) &= f(x + (y-x)) \\ &= f(y). \\ \phi(0) &= f(x + 0(y-x)) \\ &= f(x). \end{aligned}$$

Luego, por el Teorema Fudnamental del Cálculo para $\phi(\alpha)$ tenemos que

$$\begin{aligned} f(y) - f(x) &= \phi(1) - \phi(0) \\ &= \int_0^1 \phi'(\alpha) d\alpha \\ &= \int_0^1 \underbrace{\nabla^T f(x + \alpha(y-x))(y-x)}_{(8)} d\alpha \\ &\geq \int_0^1 \underbrace{\nabla^T f(x)(y-x)}_{(7)} d\alpha \\ &= \nabla^T f(x)(y-x) \int_0^1 d\alpha \\ &= \nabla^T f(x)(y-x). \end{aligned}$$

entonces

$$f(y) - f(x) \geq \nabla^T f(x)(y - x)$$

pero este último enunciado es equivalente a que $f(x)$ sea convexa. $\square$