Introduction to Optimization — Homework 1

Dimitri Tabatadze · 02:20, February 24, 2024

1. (a) Let $n = 1,\ S = [0, \infty) \subset \mathbb{R}^1,\ C = [1, \infty) \subset S,\ f(x) = x^2.\ f(x)$ attains its minumum but not the maximum.

   (b) i. Let $n = 1,\ S = (0, 2),\ C = (0, 1),\ f(x) = x.$

   ii. Let $n = 1,\ S = (0, 2),\ C = (0, 1),\ f(x) = -x(x - 2).$

2. (a) The plots:

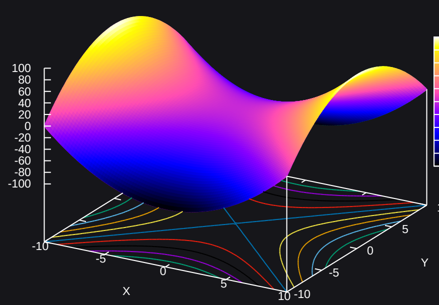

Figure 1: plot of f

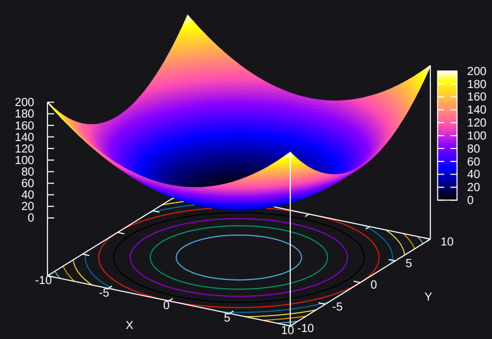

Figure 2: plot of g

   (b) Since D is not closed, f doesn't attain either a minumum or a maximum but g attains a minimum at $(0, 0)$ (very clear form the plot):

$$\nabla g(x, y) = \langle 2x, 2y \rangle,\ \nabla g(0, 0) = \langle 0, 0 \rangle$$

   (c) Now, since $\bar{D}$ is closed, out of continuity of f and g, both attain both, a minumum and a maximum:

$$\begin{aligned} &\max_{x \in \bar{D}} f(x) = 1, \quad x = \langle \pm 1, 0 \rangle &&\max_{x \in \bar{D}} g(x) = 1, \quad x = \langle \cos\alpha, \sin\alpha \rangle \forall \alpha \in [-, 2\pi] \\ &\max_{x \in \bar{D}} f(x) = -1, \quad x = \langle 0, \pm 1 \rangle &&\max_{x \in \bar{D}} g(x) = 0, \quad x = \langle 0, 0 \rangle \end{aligned}$$

3. (a) Since all eigenvalues are ≥ 0 the matrix is *positive semidefinite.*

$$\begin{aligned} \det(M - \lambda I) &= (2 - \lambda)^3 - 2 - 3(2 - \lambda) \\ &= -\sigma^3 - 2 + 3\sigma && \sigma \equiv \lambda - 2 \\ &= -(\sigma - 1)^2(\sigma + 2) && \sigma = -2, 1 \implies \boxed{\lambda = 0, 3} \\ &= 0 \end{aligned}$$

   (b) Since all eigenvalues are > 0 the matris is *positive definite*

$$\begin{aligned} \det(M - \lambda I) &= (2 - \lambda)^3 - 3(2 - \lambda) + 2 \\ &= -\sigma^3 + 3\sigma + 2 && \sigma \equiv \lambda - 2 \\ &= (\sigma + 1)^2(\sigma - 2) && \sigma = -1, 2 \implies \boxed{\lambda = 1, 4} \end{aligned}$$

(c) Since all eigenvalues are > 0 the matris is *positive definite*

$$\begin{aligned}\det(M-\lambda I) &= (1-\lambda)(5-\lambda)(9-\lambda)+48-16(1-\lambda)-9(5-\lambda)-4(9-\lambda)\\ &= -\lambda(\lambda^2-15\lambda+30)\\ &= 0 \implies \boxed{\lambda = 0, \frac{15-\sqrt{105}}{2}, \frac{15+\sqrt{105}}{2}}\end{aligned}$$

4. (a) Since f is just the distance function from Ax to b squared, we can say that when $Ax = b$ the distance will be minimized. To find such x we can write:

$$x = A^{\dagger}b = R^{-1}Q^Tb$$

(b) Let's rewrite f as such:

$$f(x_1,\ldots,x_n) = \sum_{i=1}^{m}\frac{1}{2}\left(\sum_{j=1}^{n}x_j\cdot a_{ij}-b_i\right)$$

and then

$$\begin{aligned}H_f(x) &= \left(\frac{\delta^2 f(x)}{\delta x_i \delta x_j}\right)_{i,j=1,\ldots,n}\\ &= \left(\frac{\delta}{\delta x_i}\cdot\frac{\delta f(x)}{\delta x_j}\right)_{i,j=1,\ldots,n}\\ &= \left(\frac{\delta}{\delta x_i}\cdot\left(\sum_{p=1}^{m}a_{pj}\left(\sum_{q=1}^{n}x_q\cdot a_{pq}-b_p\right)\right)\right)_{i,j=1,\ldots,n}\\ &= \left(\sum_{p=1}^{n}a_{pj}\cdot a_{pi}\right)_{i,j=1,\ldots,n}\\ &= A^TA\end{aligned}$$

$$x^T(A^TA)x = \langle Ax, Ax\rangle \geq 0 \iff A^TA \text{ positive semidefinite}$$

if A is injective, $Ax = 0 \implies x = 0$, then $\langle Ax, Ax\rangle = 0 \implies x = 0$, therefore $\langle Ax, Ax\rangle > 0\ \forall\ x \neq 0$

(c) let

$$A = \begin{pmatrix}\xi_1 & \ldots & \xi_m\\ 1 & \ldots & 1\end{pmatrix}^T,\ b = \begin{pmatrix}\eta_1 & \ldots & \eta_m\end{pmatrix}^T,\ x = \begin{pmatrix}x_1 & x_2\end{pmatrix},\ g(y) = x_1\cdot y + x_2$$

then $\bar{x}$ which minimizes f will also minimize $\sum_{i=1}^{m}(g(\xi_i)-\eta_i)^2$. Since $H_f = A^TA$ we get that

$$H_f = \begin{pmatrix}\sum\limits_{i=1}^{m}\xi_i^2 & \sum\limits_{i=1}^{m}\xi_i\\ \sum\limits_{i=1}^{m}\xi_i & m\end{pmatrix}.$$

Proof. Since A^TA is always positive semidefinite, what's left to show is that it's invertilbe only when there are at ldeast one pair of $i \neq j$ where $\xi_i \neq \xi_j$.

$$\begin{aligned}
\det(H_f) &= m\sum_{i=1}^{m}\xi_i^2 - \left(\sum_{i=1}^{m}\xi_i\right)^2 \\
&= \sum_{j=1}^{m}\sum_{i=1}^{m}\xi_i^2 - \sum_{j=1}^{m}\xi_j\sum_{i=1}^{m}\xi_i \\
&= \frac{1}{2}\sum_{j=1}^{m}\sum_{i=1}^{m}\xi_i^2 + \frac{1}{2}\sum_{j=1}^{m}\sum_{i=1}^{m}\xi_j^2 - \sum_{j=1}^{m}\sum_{i=1}^{m}\xi_i\xi_j \\
&= \frac{1}{2}\sum_{j=1}^{m}\sum_{i=1}^{m}\left(\xi_i^2 + \xi_j^2 - 2\xi_i\xi_j\right) \\
&= \frac{1}{2}\sum_{j=1}^{m}\sum_{i=1}^{m}(\xi_i - \xi_j)^2
\end{aligned}$$

Since $(\xi_i - \xi_j)^2$ is always ≥ 0, $\det(H_f) = 0 \iff (\xi_i - \xi_j)^2 = 0 \ \forall\ i,j \iff \xi_i = \xi_j \ \forall\ i,j$[1]. □

5. (a) By definition of totaly differentiability

$$f(x+d) = f(x) + \nabla f(x)^T d + o(\|d\|) \quad \text{for } \|d\| \to 0$$
$$\Downarrow$$
$$f(x+d) - f(x) = \nabla f(x)^T d + o(\|d\|) \quad \text{for } \|d\| \to 0$$

(b) By Taylor expansion if $f : \mathbb{R}^n \to \mathbb{R}$ is two times continiuously differentiable

$$f(x+h) = f(x) + \nabla f(x)^T h + \frac{1}{2}h^T H_f(x)h + o(\|h\|^2)$$
$$\Downarrow$$
$$f(x+h) - f(x) = \nabla f(x)^T h + \frac{1}{2}h^T H_f(x)h + o(\|h\|^2)$$
$$\Downarrow$$
$$f(x+h) - f(x) = \nabla f(x)^T h + O(\|h\|^2)$$

(c) By Taylor expansion if $f : \mathbb{R}^n \to \mathbb{R}$ is two times continiuously differentiable

$$f(x+h) = f(x) + \nabla f(x)^T h + \frac{1}{2}h^T H_f(x)h + o(\|h\|^2)$$
$$\Longrightarrow$$
$$f(x+h) - f(x) = \nabla f(x)^T h + \frac{1}{2}h^T H_f(x)h + o(\|h\|^2)$$

6. (a) Compatibility:

$$\begin{aligned}
& \|Ay\| \leq \|A\| \cdot \|y\| & \\
\Longrightarrow \quad & \|A\| \geq \frac{\|Ay\|}{\|y\|} & \text{for } \|y\| \neq 0 \\
\Longrightarrow \quad & \|A\| \geq \left\|A\frac{y}{\|y\|}\right\| & \text{for } \|y\| \neq 0 \\
\Longrightarrow \quad & \|A\| \geq \|Ay\| & \text{for } \|y\| = 1 \\
\Longrightarrow \quad & \|A\| = \max_{\|y\|=1}\|Ay\| &
\end{aligned}$$

(b) Sub-multiplicativity[2]:

$$\begin{aligned}
||AB|| &= \max_{x\neq 0}\frac{||ABx||}{||x||} \\
&= \max_{Bx\neq 0}\frac{||ABx||}{||x||} \\
&= \max_{Bx\neq 0}\frac{||ABx||}{||Bx||}\frac{||Bx||}{||x||} \\
&\leq \max_{y\neq 0}\frac{||Ay||}{||y||}\max_{x\neq 0}\frac{||Bx||}{||x||}
\end{aligned}$$

[1] a friend helped me with this one.
[2] https://math.stackexchange.com/questions/435621/show-that-the-operator-norm-is-submultiplicative