



Introduction to Optimization — Homework 1

Dimitri Tabatadze · 12:21, February 27, 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 minimum 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:

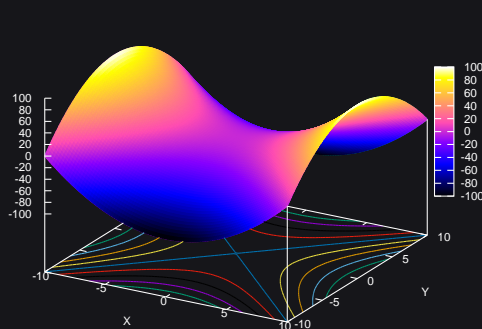


Figure 1: plot of f

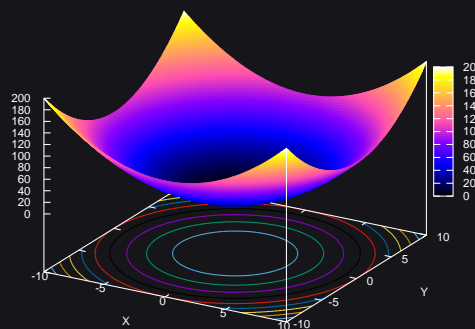


Figure 2: plot of g

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

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

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

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

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 matrix 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 matrix 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:

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

(b) Let's rewrite f as such:

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

and then

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

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

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

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

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

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

Proof. ¹ Since $A^T A$ is always positive semi-definite, what's left to show is that it's invertible only when

¹a friend helped me with this one.

there are at least one pair of $i \neq j$ where $\xi_i \neq \xi_j$.

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

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

5. (a) By definition of total differentiability

$$\begin{aligned}
f(x+d) &= f(x) + \nabla f(x)^T d + o(\|d\|) \quad \text{for } \|d\| \rightarrow 0 \\
&\Downarrow \\
f(x+d) - f(x) &= \nabla f(x)^T d + o(\|d\|) \quad \text{for } \|d\| \rightarrow 0
\end{aligned}$$

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

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

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

$$\begin{aligned}
f(x+h) &= f(x) + \nabla f(x)^T h + \frac{1}{2} h^T H_f(x) h + o(\|h\|^2) \\
&\implies \\
f(x+h) - f(x) &= \nabla f(x)^T h + \frac{1}{2} h^T H_f(x) h + o(\|h\|^2)
\end{aligned}$$

6. (a) Compatibility:

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

(b) Sub-multiplicativity²:

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

²<https://math.stackexchange.com/questions/435621/show-that-the-operator-norm-is-submultiplicative>