046197

Computational Methods in Optimization

HOMEWORK #1

Alexander Shender: 328626114

Eliran Cohen: 204187801

Contents

Question 1.

Part a.

There are 4 conditions that need to hold that define the inner product, as specified in the tutorial no. 1:

1. סימטריה: $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$.
2. אדיטיביות: $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$.
3. הומוגניות: $\langle \lambda \mathbf{x}, \mathbf{y} \rangle = \lambda \langle \mathbf{x}, \mathbf{y} \rangle$.
4. חיוביות מוגדרת: $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ וכן $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ אם ורק אם $\mathbf{x} = \mathbf{0}$.

In our case, we have $\langle x, y \rangle_M = x^T My$; $M \ is \ a \ general \ matrix \in R^{n \times n}$

Condition 1 holds:

$$\langle x, y \rangle_M = x^T My \underbrace{=}_{it's\ a\ scalar} (x^T My)^T = y^T Mx = \langle y, x \rangle_M$$

Condition 2 holds:

$$\langle x + y, z \rangle_M = (x + y)^T Mz = x^T Mz + y^T Mz = \langle x, z \rangle_M + \langle y, z \rangle_M$$

Condition 3 holds:

$$\langle \lambda x, y \rangle_M = \lambda x^T My = \lambda \langle x, y \rangle_M$$

While conditions 1-3 indeed hold, the condition 4 will hold only if $x^T Mx \geq 0$ and $x \neq 0$. This means M matrix should at least PSD. But we don't have such condition on M.

As basic example let's take M to be a matrix of all zeros. Then $\langle x, x \rangle_M = x^T Mx = 0$, but $x \neq 0$.

So this operation is **NOT AN INNER PRODUCT**

Part b.

The only difference we have here is the new information on the matrix (now denoted Q): $Q \succ 0$

So conditions 1-3 hold as previously in Part a.

Condition 4:

$$\langle x, x\rangle_Q = x^T Q x \geq 0 \quad ; \quad \langle x, x\rangle_Q \; iff \; x = 0$$

Since we know that Q is a PD matrix, this means that for each $x \in R^n$; $x \neq 0$,

$$x^T Q x > 0$$

Which helps us to prove the condition 4.

So this operation **answers all 4 conditions for the inner product.**

Question 2.

Given: A is symmetric, $A \in R^{n\times x}$

Part a.

Given: $A \succ 0$. Prove: A is invertible, $A^{-1} \succ 0$

Invertibility: If A is symmetric, PD, $A \in R^{n\times n}$, then the following hold from the characteristics of the PD matrixes:

$$x^T Ax > 0 \quad \forall x \neq 0$$

Which means:

$$Ax > 0 \quad \forall x \neq 0$$

Thus, A spans over all of its dimensions and has a full rank. **Thus, A is invertible**.

Another way to show this is to remember that all of eigenvalues of A are positive -> A is invertible.

To prove that A^{-1} is PD, we show that all its eigenvalues are strictly positive:

For matrix A, eigenvector $v \neq 0$ and corresponding eigenvalue $\lambda > 0$:

$$Av = \lambda v$$

$$A^{-1}Av = A^{-1}\lambda v \quad ; \quad A^{-1}A = I$$

$$\frac{1}{\lambda}v = A^{-1}v$$

Since $\lambda > 0$, so does $\frac{1}{\lambda} > 0$. Also, $v \neq 0$.

Thus, all the eigenvalues are positive, and $\boldsymbol{A^{-1} \succ 0}$.

Part b.

Given: $A \succcurlyeq 0$, A invertible. Prove: $A \succ 0$

A is invertible -> $\lambda \neq 0 \quad \forall \lambda \in (eigenvalues\ of\ A)$

Invertible matrix has all non-zero eigenvalues.

$A \succcurlyeq 0$ -> $\lambda \geq 0 \quad \forall \lambda \in (eigenvalues\ of\ A)$

Combining two of those constraints together we get:

$$\lambda > 0 \quad \forall \lambda \in (eigenvalues\ of\ A)$$

Which is exactly the property of PD matrix – all its eigenvalues are strictly positive.

Thus, $A \succ 0$

Part c.

Given: $A = B^T B$; $B \in R^{m \times n}$. Prove: $A \succcurlyeq 0$

We try to prove the property of the PDS matrix:

$$x^T A x = x^T B^T B x \underbrace{=}_{C=Bx} C^T C = \|C\|_2^2 \geq 0$$

Which means

$$x^T A x \geq 0$$

Thus, $A \succcurlyeq 0$.

Example for B: m = 1, n = 2.

$$B \neq 0 = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

$$A = B^T B = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

We take $x \neq 0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$:

$$x^T A x = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0$$

Part d.

Given: $A_{i,i} < 0 \quad ; \quad 1 \leq i \leq n,$ Prove: $A \ngeq 0$.

It if sufficient to show that for some $x \neq 0$, this holds: $x^T Ax < 0$

Expanding:

$$x^T Ax = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j A_{i,j} = \sum_{i=1}^{n} x_i^2 A_{i,i} + \sum_{i=1}^{n} \sum_{j \neq i}^{n} x_i x_j A_{i,j}$$

If some $A_{i,i} < 0$, this means this whole expression can be negative too, if we choose $x \neq 0$, that has all the values equal 0 except for the $i'th$ value. For example:

$$x^T Ax = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & 2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -1 \ngeq 0$$

The condition that A has non-negative diagonal values is **sufficient** to prove it is PSD, Example for the extreme case by having zero values on the diagonals:

$$x^T Ax = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 0 \geq 0$$

By having positive values on a diagonal will still make the inequality hold.

Question 3.

Given:

$f : U \to R$ continuously differentiable. $U \in R^n$. $d \neq 0$ is the downhill direction at point x if there exists T such that for each $0 < t < T$:

$$f(x + td) < f(x)$$

Part a.

Prove: If direction d uphold: $f'(x, d) < 0$ for some x, then d is the downhill direction.

From the definition:

$$f'(x, d) = \lim_{t \to 0^+} \frac{f(x + td) - f(x)}{t}$$

$$\lim_{t \to 0^+} \frac{f(x + td) - f(x)}{t} < 0$$

$$f(x + td) - f(x) < 0$$

$$f(x + td) < f(x)$$

Thus, this is the downhill direction from the definition given in this exercise.

Part b.

Prove: if $\nabla f(x) \neq 0$, then the direction $d = -\nabla f(x)$ is the downhill direction.

From theorem 4 from the tutorial: if f is continuously differentiable, then:

$$f'(x; d) = \nabla f(x)^T d$$

If $d = -\nabla f(x)$, then $\nabla f(x) = -d$, then:

$$f'(x; d) = \nabla f(x)^T d = -d \cdot d = -d^2 < 0$$

Thus,

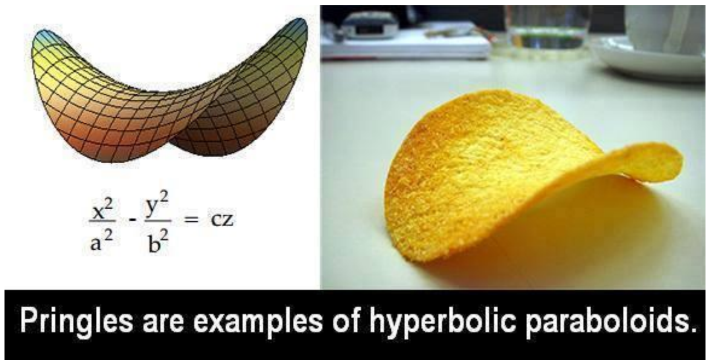
$$f'(x, d) < 0$$

From Part a. of this question we know that if this inequality holds, then d is the downhill direction.

Part c.

Find example of a function $f(x_1, x_2)$, a point $x^0 = [x_1^0\ x_2^0]^T$, directions d_1, d_2, s.t. d_1 is the downhill direction and d_2 is not.

What came on my mind is Pringles (image source – google images):



Let's take $a = 1; b = 1, c = 1$

$$f(x_1, x_2) = x_1^2 - x_2^2$$

Let's take trivial point $\begin{cases} x_1 = 1 \\ x_2 = 1 \end{cases}$

Calculating gradient:

$$\nabla f(x) = \begin{bmatrix} \frac{\delta f}{\delta x_1} & \frac{\delta f}{\delta x_2} \end{bmatrix} = [2x_1 \quad -2x_2] = [2 \quad -2]$$

From what we have proven:

$$\begin{cases} d_1 = \nabla f(x_1, x_2) = [2 \quad -2] = UPHILL\ DIRECTION \\ d_2 = -\nabla f(x_1, x_2) = [-2 \quad 2] = DOWNHILL\ \ DIRECTION \end{cases}$$

Question 4-6.

<continue to the next page>