Assignment 1

by Adil Akhmetov

Quesion 1.

- a.

Let $\boldsymbol{x} \in \mathbb{R}^M, \boldsymbol{y} \in \mathbb{R}^N$, function f: $\mathbb{R}^M \times \mathbb{R}^N \rightarrow \mathbb{R}$, $f(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^T \boldsymbol{A}\boldsymbol{y} + \boldsymbol{x}^T \boldsymbol{B}\boldsymbol{x} - \boldsymbol{C}\boldsymbol{y} + \boldsymbol{D}$. Compute the dimensions of the matrices **A**, **B**, **C**, **D** for the function so that the mathematical expression is valid.

$$\begin{gathered} x \in R^M,\ then\ R^{M \times 1} \\ y \in R^N,\ then\ R^{N \times 1} \\ f(x,y) = x^T Ay + x^T Bx - Cy + D = \\ (R^M)^T A(R^N) + (R^M)^T B(R^M) - C(R^N) + D = \\ R^{1 \times M} AR^N + R^{1 \times M} BR^M - CR^N + D = \\ R^{1 \times M} R^{M \times N} R^N + R^{1 \times M} R^{M \times M} R^M - R^{1 \times N} R^N + R^{1 \times 1} = \\ R + R - R + R = 2R \\ Thus, \\ A \in R^{M \times N},\ B \in R^{M \times M},\ C \in R^{N \times 1},\ D \in R^{1 \times 1} \end{gathered}$$

- b.

Let $\boldsymbol{x} \in \mathbb{R}^N, \boldsymbol{M} \in \mathbb{R}^{N \times N}$. Express the function $f(\boldsymbol{x}) = \sum_{i=1}^N \sum_{j=1}^N x_i x_j M_{ij}$

$$\begin{gathered} f(x) = \sum_{i=1}^N \sum_{j=1}^N x_i x_j M_{ij} = \\ \sum_{i=1}^N (\sum_{j=1}^N x_j M_{ij}) x_i = \\ \sum_{i=1}^N x^T M_i x_i = \\ x^T \sum_{i=1}^N M_i x_i = \\ x^T Mx = \\ R^{1 \times N} R^{N \times M} R^{N \times 1} = R \end{gathered}$$

P.S. I hope, that I was correct in the notation. The result must be scalar, so the eventual result must be correct, but I am not sure about the approach.

- c.

Suppose $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V}$, where $\boldsymbol{V}$ is a vector space. $||\boldsymbol{u}|| = ||\boldsymbol{v}|| = 1$ and $\boldsymbol{u}^T \boldsymbol{v} = 1$. Prove that $\boldsymbol{u} = \boldsymbol{v}$

Let's use the following formula:

$$\theta = \cos^{-1} \frac{u^T v}{|u||v|}$$

Then,

$$\theta = \cos^{-1} \frac{1}{1 * 1} = cos^{-1}(1) = 0$$

So, the angle between the vectors is 0. It means, that the vectors are directed to the one side. And since the lengths of the vectors are equal, the vectors are the same.

Quesion 2.

Let $\boldsymbol{A} \in \mathbb{R}^{M \times N}, \boldsymbol{x} \in \mathbb{R}^{M}$, and $\boldsymbol{b} \in \mathbb{R}^{M}$ that forms a system of linear equations: $\boldsymbol{Ax} = \boldsymbol{b}$

- Suppose M = N = 2, matrix $\boldsymbol{A}$ has the form of: $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$
  Derive the conditions that this system of linear equation has a unique solution for $\boldsymbol{x}$
- Suppose M = N = 10, and the matrix **A** has the following eigenvalues { 1, 3, -2, 10, 0, 4, 2, 9, 7, -1 }. Dose the system has a unique solution for $\boldsymbol{x}$?

- So, I think, that if the determinant of the matrix is 0, then the matrix is singular. And if the matrix is singular, then the matrix is not invertible. Thus, it has a linear dependency in it. Hence, that this system of linear equation has no solution. The main condition is:

$$\det(A) \neq 0$$
$$ad - bc \neq 0$$

- The set of eigenvalues contains 0, which means that the kernel (nullspace) of the matrix is nonzero. This means that the matrix has determinant equal to zero. Then, according to the solution above, the matrix is singular and not invertible. Hence, that this system of linear equation has no solution.

Quesion 3.

- a.

  Consider the quadratic function f: $\mathbb{R}^N \to \mathbb{R}$ with $f(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x}$ for a symmetric matrix A. Determine the gradient of the function f.

$$f(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} x_i x_j$$
$$\frac{\delta f(x)}{\delta x_k} = \frac{\delta}{\delta x_k} \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} x_i x_j =$$
$$\frac{\delta}{\delta x_k} [\sum_{i \neq k} \sum_{j \neq k} A_{ij} x_i x_j + \sum_{i \neq k} A_{ik} x_i x_k + \sum_{j \neq k} A_{kj} x_k x_j + A_{kk} x_k^2] =$$
$$\sum_{i \neq k} A_{ik} x_i + \sum_{j \neq k} A_{kj} x_j + 2A_{kk} x_k =$$
$$\sum_{i=1}^{n} A_{ik} x_i + \sum_{j=1}^{n} A_{kj} = 2 \sum_{i=1}^{n} A_{ki} x_i$$
$$\frac{\delta^2 f(x)}{\delta x_k \delta x_l} = 2A_{lk} = 2A_{kl}$$
$$\nabla_x^2 f(x) = 2A$$

P.S. Taken from lecture 2 slide 67.

Question 4.

Compute the derivatives for the following functions f: $\mathbb{R} \to \mathbb{R}$

- $f: f(x) = (x^5 + x^3 + 1)^2$
- $g: g(x) = \frac{e^{9x^2+7} - 1}{e^{9x^2+7} + 1}$
- $h: h(x) = (1 - x)\log(1 - x)$

- $$f'(x) = ((x^5 + x^3 + 1)^2)' =$$
$$2(x^5 + x^3 + 1)^{2-1} \times (x^5 + x^3 + 1)' =$$
$$2(x^5 + x^3 + 1) \times (5x^4 + 3x^2 + 0) =$$
$$10x^9 + 16x^7 + 6x^5 + 10x^4 + 6x^2$$

- $$g'(x) = (\frac{e^{9x^2+7}-1}{e^{9x^2+7}+1})' =$$
$$\frac{(e^{9x^2+7}-1)'(e^{9x^2+7}+1)-(e^{9x^2+7}-1)(e^{9x^2+7}+1)'}{(e^{9x^2+7}+1)^2} =$$
$$\frac{((e^{9x^2+7})'-0)(e^{9x^2+7}+1)-(e^{9x^2+7}-1)((e^{9x^2+7})'+0)}{(e^{9x^2+7}+1)^2} =$$
$$\frac{(e^{9x^2+7}(9x^2+7)'-0)(e^{9x^2+7}+1)-(e^{9x^2+7}-1)(e^{9x^2+7}(9x^2+7)'+0)}{(e^{9x^2+7}+1)^2} =$$
$$\frac{(e^{9x^2+7}(18x+0)-0)(e^{9x^2+7}+1)-(e^{9x^2+7}-1)(e^{9x^2+7}(18x+0)+0)}{(e^{9x^2+7}+1)^2} =$$
$$\frac{(e^{9x^2+7}(18x+0)-0)(e^{9x^2+7}+1)-(e^{9x^2+7}-1)(e^{9x^2+7}(18x+0)+0)}{(e^{9x^2+7}+1)^2} =$$
$$\frac{18xe^{9x^2+7}(e^{9x^2+7}+1)-18xe^{9x^2+7}(e^{9x^2+7}-1)}{(e^{9x^2+7}+1)^2} =$$
$$\frac{36xe^{9x^2+7}}{(e^{9x^2+7}+1)^2}$$

- $$h'(x) = ((1-x)log(1-x))' =$$
$$(1-x)'log(1-x)+(log(1-x))'(1-x) =$$
$$(0-1)log(1-x)+\frac{(1-x)'}{1-x}(1-x) =$$
$$-log(1-x)+(0-1) =$$
$$-log(1-x)-1$$

Question 5.

Calculate the gradients of the following functions. $f\colon \mathbb{R}^2 \to \mathbb{R}$,

- $f\colon f(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{x}\|_2^2$
- $f\colon f(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{x}\|_2$

- $$f(x) = \frac{1}{2}||x||_2^2 = \frac{1}{2}\left(\sum_{i=1}^{n}|x_i|^2\right)$$
$$f'(x) = \frac{\delta f(x)}{\delta x_k} = \frac{1}{2}\frac{\delta}{\delta x_k}\left[\sum_{i\neq k}^{n}|x_i|^2+|x_k|^2\right] =$$
$$\frac{1}{2}\left[0+2|x_k|^{2-1}\right] = |x_k|$$

- $$f(x) = \frac{1}{2}||x||_2 = \frac{1}{2}\left(\sum_{i=1}^{n}|x_i|^2\right)^{\frac{1}{2}}$$
$$f'(x) = \frac{\delta f(x)}{\delta x_k} = \frac{1}{2}\frac{\delta}{\delta x_k}\left[\left(\sum_{i\neq k}^{n}|x_i|^2+|x_k|^2\right)^{\frac{1}{2}}\right] =$$
$$\frac{1}{2}\left[\frac{1}{2}\left(0+2|x_k|^{2-1}\right)^{\frac{1}{2}-1}\right] = \frac{1}{4\sqrt{2|x_k|}}$$

P.S. I'm not sure this is correct, I couldn't figure out how to properly calculate the gradient, so I used the solution from Question 3 as an example

Question 6.

Let $\mathbf{A} = \begin{bmatrix} 0 & 2 & 4 \\ 2 & 4 & 2 \\ 3 & 3 & 1 \end{bmatrix}$, $\boldsymbol{b} = [-2 \quad -2 \quad -4]^T$, $\boldsymbol{c} = [1 \quad 1 \quad 1]^T$

a. What is $\mathbf{Ac}$?

b. What is the solution to the linear system $\mathbf{A}\boldsymbol{x} = \mathbf{b}$

- a.

$$Ac = \begin{bmatrix} 0 & 2 & 4 \\ 2 & 4 & 2 \\ 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T =$$
$$\begin{bmatrix} 0*1+2*1+4*1 \\ 2*1+4*1+2*1 \\ 3*1+3*1+1*1 \end{bmatrix} =$$
$$\begin{bmatrix} 0+2+4 \\ 2+4+2 \\ 3+3+1 \end{bmatrix} = \begin{bmatrix} 6 \\ 8 \\ 7 \end{bmatrix}$$

- c.

$$Ax = b$$
$$x = A^{-1}b$$
$$det(A) = \begin{vmatrix} 0 & 2 & 4 \\ 2 & 4 & 2 \\ 3 & 3 & 1 \end{vmatrix} =$$
$$-\begin{vmatrix} 2 & 4 & 2 \\ 0 & 2 & 4 \\ 3 & 3 & 1 \end{vmatrix} =$$
$$-\begin{vmatrix} 2 & 4 & 2 \\ 0 & 2 & 4 \\ 0 & -3 & -2 \end{vmatrix} =$$
$$-\begin{vmatrix} 2 & 4 & 2 \\ 0 & 2 & 4 \\ 0 & 0 & 4 \end{vmatrix} =$$
$$-(2*2*4) = -16$$
$$det(A) \neq 0$$
$$Then,$$
$$A^{-1} = \begin{bmatrix} \frac{1}{8} & -\frac{5}{8} & \frac{3}{4} \\ -\frac{1}{4} & \frac{3}{4} & -\frac{1}{2} \\ \frac{3}{8} & -\frac{3}{8} & \frac{1}{4} \end{bmatrix}$$
$$x = \begin{bmatrix} \frac{1}{8} & -\frac{5}{8} & \frac{3}{4} \\ -\frac{1}{4} & \frac{3}{4} & -\frac{1}{2} \\ \frac{3}{8} & -\frac{3}{8} & \frac{1}{4} \end{bmatrix} \times \begin{bmatrix} -2 \\ -2 \\ -4 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}$$

Question 7.

- a.

Let $\mathbf{A} \in \mathbb{R}^{M \times N}$ be orthonormal and $\boldsymbol{x} \in \mathbb{R}^N$. An orthonormal matrix is a square martix whose columns and rows are orthonormal vectors, such that $\mathbf{AA}^\mathrm{T} = \mathbf{A}^\mathrm{T}\mathbf{A} = \mathbf{I}$ where $\mathbf{I}$ is the identity matrix. Show that $\|\mathbf{A}\boldsymbol{x}\|_2^2 = \|\boldsymbol{x}\|_2^2$

First, I want to note and if $M \neq N$, then $AA^T \neq I$ according to the lecture notes. But, maybe I'm wrong

$$||Ax||_2^2 = (Ax)^T Ax = A^T x^T Ax = A^T A x^T x = I x^T x = x^T x = ||x||_2^2$$

- b.

  Let $\mathbf{B} \in \mathbb{R}^{M\times N}$ be invertible and symmetric. A symmetric matrix is a square matrix satisfying $\mathbf{B} = \mathbf{B}^{\mathrm{T}}$. Show that $\mathbf{B}^{-1}$ is also symmetric.

Same here, I'm not sure that if $M \neq N$ A matrix may be symmetric.

$$\begin{gathered} Since,\ I = I^T\ and\ AA^{-1} = I \\ AA^{-1} = (AA^{-1})^T \\ AA^{-1} = (A^{-1})^T A^T \\ AA^{-1} = (A^{-1})^T A \\ A^{-1}AA^{-1} = (A^{-1})^T AA^{-1} \\ A^{-1}I = (A^{-1})^T I \\ A^{-1} = (A^{-1})^T \end{gathered}$$

- c.

  Let $\mathbf{C} \in \mathbb{R}^{M\times N}$ be positive semi-definite (PSD). A positive semi-definite matrix is a symmetric matrix satisfying $\boldsymbol{x}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\boldsymbol{x} \geq 0$ for any vector $\boldsymbol{x} \in \mathbb{R}^N$. Show that its eigenvalues are non-negative.

  I guess, there is a typo in the condition of PSD. It must've been $x^T Cx \geq 0$

$$\begin{gathered} Ax = \lambda x \\ Suppose \\ x = \vec{v}\ (an\ eigenvector) \\ Then, \\ A\vec{v} = \lambda\vec{v} \\ \vec{v}^T A\vec{v} \geq 0 \\ \vec{v}^T \lambda\vec{v} \geq 0 \end{gathered}$$

P.S. I didn't figure out how to prove it, maybe if eigenvectors are nonnegative, then λ must be nonnegative as well?

Question 8.

The trace of a matrix is the sum of the diagonal entries; $Tr(\mathbf{A}) = \sum_i \mathbf{A}_{ii}$. If $\mathbf{A} \in \mathbb{R}^{M\times N}$ and $\mathbf{B} \in \mathbb{R}^{N\times M}$, show that $Tr(\mathbf{AB}) = Tr(\mathbf{BA})$.

$$Tr(AB) = \sum_{i=1}^{m}(AB)_{ii} = \sum_{i=1}^{m}(\sum_{j=1}^{n} A_{ij}B_{ji}) = \sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}B_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{m} B_{ji}A_{ij} = \sum_{j=1}^{n}(\sum_{i=1}^{m} B_{ji}A_{ij}) = \sum_{j=1}^{n}(BA)_{jj} = Tr(BA)$$

P.S. taken from the lecture notes page 9

Question 9.

Compute the length (Euclidian norm) of the following vectors

$$\left\|\begin{bmatrix}0\\18\end{bmatrix}\right\|_2 = \qquad \left\|\begin{bmatrix}6\\-3\\6\end{bmatrix}\right\|_2 = \qquad \left\|\begin{bmatrix}\sqrt{35}\\1\end{bmatrix}\right\|_2 =$$

- $$\left|\left|\begin{bmatrix}0\\18\end{bmatrix}\right|\right|_2 = (|0|^2 + |18|^2)^{\frac{1}{2}} = 18$$
- $$\left|\left|\begin{bmatrix}6\\-3\\6\end{bmatrix}\right|\right|_2 = (|6|^2 + |-3|^2 + |6|^2)^{\frac{1}{2}} = 9$$
- $$\left|\left|\begin{bmatrix}\sqrt{35}\\1\end{bmatrix}\right|\right|_2 = (|\sqrt{35}|^2 + |1|^2)^{\frac{1}{2}} = 6$$

Question 10.

Compute $\boldsymbol{u}^T\boldsymbol{v}$ and $\boldsymbol{u}^T\boldsymbol{w}$ and $\boldsymbol{u}^T(\boldsymbol{v}+\boldsymbol{w})$ and $\boldsymbol{w}^T\boldsymbol{v}$

$\boldsymbol{u} = \begin{bmatrix}-9\\12\end{bmatrix}$ $\boldsymbol{v} = \begin{bmatrix}3\\4\end{bmatrix}$ $\boldsymbol{w} = \begin{bmatrix}8\\6\end{bmatrix}$

Calculate the lengths $\|\boldsymbol{u}\|$, $\|\boldsymbol{v}\|$, and $\|\boldsymbol{w}\|$ of those vectors. Find unit vectors in the direction of **v** and **w**.

$$u^Tv = \begin{bmatrix}-9 & 12\end{bmatrix}\begin{bmatrix}3\\4\end{bmatrix} = -9*3+12*4 = 21$$
$$u^Tw = \begin{bmatrix}-9 & 12\end{bmatrix}\begin{bmatrix}8\\6\end{bmatrix} = -9*8+12*6 = 0$$
$$w^Tv = \begin{bmatrix}8 & 6\end{bmatrix}\begin{bmatrix}3\\4\end{bmatrix} = 8*3+6*4 = 48$$
$$||u|| = \sqrt{(-9)^2+(12)^2} = \sqrt{225} = 15$$
$$||v|| = \sqrt{3^2+4^2} = 5$$
$$||w|| = \sqrt{8^2+6^2} = 10$$
$$\vec{v} = \frac{1}{5}\begin{bmatrix}3\\4\end{bmatrix} = \begin{bmatrix}\frac{3}{5}\\\frac{4}{5}\end{bmatrix}$$
$$\vec{w} = \frac{1}{10}\begin{bmatrix}8\\6\end{bmatrix} = \begin{bmatrix}\frac{8}{10}\\\frac{6}{10}\end{bmatrix}$$

Question 11.

For which numbers b is this matrix positive definite? $\boldsymbol{A} = \begin{bmatrix}9 & b\\b & 16\end{bmatrix}$

I'll use determinants test, so that $1 > 0$ and $det(A) > 0$

$$det(A) > 0$$
$$9*16 - b*b > 0$$
$$144 - b^2 > 0$$
$$144 > b^2$$
$$b^2 < 144$$
$$b < \sqrt{144}$$
$$b < 12$$
$$b > -12$$
$$b \in (-12, 12)$$

Question 12.

Factor this matrix into $\boldsymbol{A} = \boldsymbol{S\Lambda S^{-1}}$. Note that $\boldsymbol{A} = \begin{bmatrix}1 & 2\\0 & 3\end{bmatrix}$ and $\boldsymbol{\Lambda}$ is a diagonal matrix.

1. Find eigenvalues of A

$$A - \lambda I = \begin{bmatrix}1 & 2\\0 & 3\end{bmatrix} - \begin{bmatrix}\lambda & 0\\0 & \lambda\end{bmatrix} = \begin{bmatrix}1-\lambda & 2\\0 & 3-\lambda\end{bmatrix}$$
$$det\left(\begin{bmatrix}1-\lambda & 2\\0 & 3-\lambda\end{bmatrix}\right) = 0$$
$$(1-\lambda)(3-\lambda) - 0*2 = 0$$
$$\lambda = 1,\ \lambda = 3$$

2. Find an eigenvector associated with each eigenvalue of A

$$Eigenvector\,for\lambda = 1$$
$$A - \lambda I = \begin{bmatrix}1-1 & 2\\0 & 3-1\end{bmatrix} = \begin{bmatrix}0 & 2\\0 & 2\end{bmatrix}$$
$$By\ echelon\ method$$
$$\begin{bmatrix}1\\0\end{bmatrix} for\lambda = 1$$

$$Eigenvector\,for\lambda = 3$$

$$A - \lambda I = \begin{bmatrix} 1-3 & 2 \\ 0 & 3-3 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 0 & 0 \end{bmatrix}$$
$$By\ echelon\ method$$
$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} for \lambda = 3$$

3. Construct matrices S and Λ so that Λ is diagonal.

$$S = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$
$$\Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$
$$S^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$