COS 598 ML

S2021

HW 1

By: Nicholas Soucy

For: Dr. Salimeh

Due: 5:00 PM 02/12/2021

Q1) Linear Algebra

(a) Show that if U is an orthogonal matrix, then for all $x \in \mathbb{R}^d$, $\|x\| = \|Ux\|$ where the norm is the Euclidean norm.

Lets start with $\|Ux\|$. By the definition of Euclidean norm, $\|Ux\| = \sqrt{(Ux, Ux)}$ then we can multiply by U^T to get $\|Ux\| = \sqrt{(U^TUx, U^TUx)}$ once U is an orthogonal matrix, $U^TU = I$ so we have $\|Ux\| = \sqrt{(x,x)}$, then by the definition of Euclidean norm we then have $\boxed{\|Ux\| = \|x\|}$

QED

(b) Show that all 2x2 orthogonal matrices have the form $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ or $\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$ for some θ. Give a Geometric interpretation.

Lets start with an orthogonal matrix $X = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ where a,b,c,d are real. By the def of orthogonal $X^TX = I$ so we have

$X^TX = \begin{bmatrix} a & c \\ b & d \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ we can solve this system to get $a^2 = 1-c^2$ $ab = cd$ $cd = -ab$ and $d^2 = 1-b^2 \Rightarrow$

$\Rightarrow a^2b^2 = c^2d^2$ & $a^2 = 1-c^2 \Rightarrow (1-c^2)b^2 = c^2d^2$ & $d^2 = 1-b^2 \Rightarrow$

$\Rightarrow (1-c^2)b^2 = c^2(1-b^2)$ this implies $\underline{b = \pm c}$ then we can do the same thing with a & d to get $\underline{d = \pm a}$. Therefore

$\boxed{A = \begin{pmatrix} a & c \\ c & -a \end{pmatrix} \text{ or } \begin{pmatrix} a & -c \\ c & a \end{pmatrix}}$ let $a = \cos\theta$ & $c = \sin\theta \Rightarrow$ $\boxed{A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \text{ \& } \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}}$

This geometric interpretation of these matrices, is that they are rotations and reflections.

Q2) Probability

Let random variables X & Y be jointly distributed random variables. Further assume they are jointly continuous with joint pdf $P(x,y)$. Show ...

(i) $V[X] = E[X^2] - [E[X]]^2$, where $V[X]$ is the variance of random variable X

Lets use the definition $V[X] = Cov(X,X)$. Then we can use the def of cov to get $Cov(X,X) = V[X] = E[(X-E[X])^2]$

Then we can use algebra to expand the square.

$E[(X-E[X])^2] = E[X^2 - 2XE[X] + E[X]^2] \Rightarrow$

$\Rightarrow E[X^2] - 2E[X]^2 + E[X]^2 \Rightarrow$ $V[X] = E[X^2] - E[X]^2$

$E[E[X]] = E[X]$

(ii) If X and Y are independent then $E[(X+1)Y] = E[Y](E[X]+1)$

We start with $E[(X+1)Y]$. Once X & Y are independent random variables, we can use the theorm $E[XY] = E[X]E[Y]$, to get: $E[(X+1)Y] = E[Y]E[X+1]$

Once $E[X]$ is linear, $E[X+Y] = E[X] + E[Y]$ so we get

$E[(X+1)Y] = E[Y](E[X] + E[1])$. Then finally, $E[c]$, where $c \in [-\infty, +\infty] = c$, so we have:

$E[(X+1)Y] = E[Y](E[X]+1)$

Q2 cont) Probability

(iii) If X and Y take values in $\{0,1\}$ and covariance X & Y is zero, $\mathrm{COV}(X,Y)=0$, then X and Y are independent

X & Y are indicator functions so lets define them

$$X := \begin{cases} 1 & \text{w/ Prob } p \\ 0 & \text{w/ Prob } (1-p) \end{cases} \qquad y := \begin{cases} 1 & \text{w/ prob } q \\ 0 & \text{w/ prob } (1-q) \end{cases}$$

Once $\mathrm{Cov}(X,Y)=0$ we get $\mathrm{Cov}(X,Y) = E[XY] - E[X]E[Y] = 0$
$\Rightarrow E[XY] = E[X]E[Y]$, once X & Y are indicator functions,
$E[X] = p$ & $E[Y] = q$ so we have $E[XY] = pq$
Therefore, once $E[XY] = pq$, $E[XY] = P(X=1)$ and $P(Y=1)$
So we have $P(X=1) \cdot P(Y=1)$. By the definition of independence for events, X and Y must be independent.

Q3 | Positive (semi-)definite matrices

Using the spectral decomposition, show that

(a) A is PSD iff $\lambda_i \geq 0$ for each i.

Lets start with the spectral decomposition of a real symmetric matrix A: $A = U \Lambda U^T$ where U is a $d \times d$ orthogonal matrix, and $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_d)$. We then can algebraically get $AU = U\Lambda$, then $A\vec{u}_i = \lambda \vec{u}_i$.

For a matrix to be positive semi-definite, $\vec{x}^T A \vec{x} \geq 0$ for all $\vec{x}$.

Once $\vec{u}_i$ is a eigenvector of A, then:

$\vec{u}^T A \vec{u} = \vec{u}^T(\lambda \vec{u})$ by substitution, then we can pull out the scaler λ to get

$\vec{u}^T A \vec{u} = \vec{u}^T(\lambda \vec{u}) = \vec{u}^T \vec{u} \lambda$.

Since $\vec{u}^T \vec{u}$ has to be a positive number for $\vec{u}^T A \vec{u}$ to be greater than or equal to 0, λ must be greater than or equal to 0.

therefore, once λ must be greater than or equal to 0 for A to be a PSD, than the statement is true.

QED

Q3| Positive (semi-) definite matrices cont.

(b) A is PD iff $\lambda_i > 0$ for each i.

Lets start with the spectral decomposition of a real symmetric matrix A: $A = U^T \Lambda U$ where U is a $d \times d$ orthogonal matrix, and $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_d)$. We can then algebraiclly get $AU = U\Lambda$, then $A\vec{u}_i = \lambda \vec{u}_i$, then...

If x is any nonzero vector, then $y = Ux \neq 0$ and

$x^T A x = x^T (U^T \Lambda U) x$ by $A = U^T \Lambda U$ substitution.

Then $(x^T U^T) \Lambda (Ux) = x^T (U^T \Lambda U)$ by matrix multipicaton.

by the def of y, we then have $(x^T U^T) \Lambda (Ux) = y^T \Lambda y$

then we can use the identity from the homework to

get $y^T \Lambda y = \sum_{i=1}^{d} \lambda_i y_i^T y_i$. Once y is non zero,

$\sum_{i=1}^{d} \lambda_i y_i y_i^T > 0$, and therefore A has positive

eigenvalues, proving that A is PD iff $\lambda_i > 0$ for each i.

QED.