AVD613 - Machine Learning for Signal Processing

Tutorial 2 Solutions

Submitted by

Vatsalya Gupta
SC19B098
B.Tech. ECE VII Semester

Department of Avionics
Indian Institute of Space Science and Technology
Thiruvananthapuram - 695 547

August 2022

Solution 1

(a) Let A be denote the un-transformed parallelepiped, f be the function defining it, and T be the transformation function. Also, $J(x)$ denotes the Jacobian matrix for the transformation at x. Since, T is linear, $J(x) = T$.

$$\text{Volume of un-transformed parallelepiped} = \int_A f(x)\, dx$$

$$\text{Linear transformation} \implies \int_{T(A)} f(x)\, dx = \int_A f(T(x))|det(J(x))|\, dx$$

We can write the volume of transformed parallelepiped as

$$\int_{T(A)} dx = \int_A |det(J(x))|\, dx = \int_A |det(J(x))|\, dx = \int_A |det(T)|\, dx$$

$$= |det(T)| \int_A dx = |det(T)| \times \text{un-transformed volume}$$

$$|det(T)| = abs\left(\begin{vmatrix} 6 & 1 & 1 \\ 4 & 2 & 5 \\ 2 & 8 & 7 \end{vmatrix}\right) = |-146| = 146 \tag{1}$$

So, the volume of the parallelepiped in transformed space will be $146 \times 100 = 14600$ cubic units.

(b) Again, transformed volume $= |det(T)| \times$un-transformed volume

$$|det(T)| = abs\left(\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}\right) = 0 \tag{2}$$

So, the volume of the parallelepiped in transformed space will be $0 \times 100 = 0$ cubic units.

(c) Zero (0) volume signifies that the sides of the parallelepiped representing the length, width and height are parallel to each other in the transformed space.

Solution 2

(a) We want to find A, the vector representation of $\begin{pmatrix} 2 & -1 & 6 \end{pmatrix}^T$ in the standard basis. So, we need to find the inverse of the transformation.

$$A = T^{-1}\begin{pmatrix} 2 & -1 & 6 \end{pmatrix}^T = \begin{bmatrix} 4 & -3 & 0 \\ 2 & -1 & 2 \\ 1 & 5 & 7 \end{bmatrix}^{-1} \begin{pmatrix} 2 & -1 & 6 \end{pmatrix}^T = \frac{adj(T)}{|T|}\begin{pmatrix} 2 & -1 & 6 \end{pmatrix}^T$$

$$\text{Now,} \quad adj(T) = \begin{bmatrix} -17 & 21 & -6 \\ -12 & 28 & -8 \\ 11 & -23 & 2 \end{bmatrix} \quad \text{and} \quad |det(T)| = |T| = -32$$

$$\implies A = \frac{-1}{32} \begin{bmatrix} -17 & 21 & -6 \\ -12 & 28 & -8 \\ 11 & -23 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix} = \begin{bmatrix} 91/32 \\ 25/8 \\ -57/32 \end{bmatrix}$$

$$\text{Hence,} \quad \begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix} = \frac{91}{32} \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix} + \frac{25}{8} \begin{bmatrix} -3 \\ -1 \\ 5 \end{bmatrix} - \frac{57}{32} \begin{bmatrix} 0 \\ 2 \\ 7 \end{bmatrix} \tag{3}$$

(b) We want to find B, the vector representation of $\begin{pmatrix} 2 & -1 & 6 \end{pmatrix}^T$ in the standard basis. So, we need to find the inverse of the transformation.

$$B = T^{-1} \begin{pmatrix} 2 & -1 & 6 \end{pmatrix}^T = \begin{bmatrix} 2 & -2 & 1 \\ 2 & 1 & -2 \\ 1 & 2 & 2 \end{bmatrix}^{-1} \begin{pmatrix} 2 & -1 & 6 \end{pmatrix}^T = \frac{adj(T)}{|T|} \begin{pmatrix} 2 & -1 & 6 \end{pmatrix}^T$$

$$\text{Now,} \quad adj(T) = \begin{bmatrix} 6 & 6 & 3 \\ -6 & 3 & 6 \\ 3 & -6 & 6 \end{bmatrix} \quad \text{and} \quad |det(T)| = |T| = 27$$

$$\implies B = \frac{1}{27} \begin{bmatrix} 6 & 6 & 3 \\ -6 & 3 & 6 \\ 3 & -6 & 6 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix} = \begin{bmatrix} 8/9 \\ 7/9 \\ 16/9 \end{bmatrix}$$

$$\text{Hence,} \quad \begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix} = \frac{8}{9} \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} + \frac{7}{9} \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} + \frac{16}{9} \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} \tag{4}$$

(c) If the basis vectors are orthonormal, then we can easily obtain the inverse of the transformation matrix by making it orthonormal as well (divide every column by its magnitude) and transposing it.

$$T_0 = \begin{bmatrix} \frac{2}{\sqrt{2^2+2^2+1^2}} & \frac{-2}{\sqrt{2^2+1^2+2^2}} & \frac{1}{\sqrt{1^2+2^2+2^2}} \\ \frac{2}{\sqrt{2^2+2^2+1^2}} & \frac{1}{\sqrt{2^2+1^2+2^2}} & \frac{-2}{\sqrt{1^2+2^2+2^2}} \\ \frac{1}{\sqrt{2^2+2^2+1^2}} & \frac{2}{\sqrt{2^2+1^2+2^2}} & \frac{2}{\sqrt{1^2+2^2+2^2}} \end{bmatrix} = \begin{bmatrix} 2/9 & -2/9 & 1/9 \\ 2/9 & 1/9 & -2/9 \\ 1/9 & 2/9 & 2/9 \end{bmatrix}$$

$$\implies T^{-1} = T_0^T = \begin{bmatrix} 2/9 & 2/9 & 1/9 \\ -2/9 & 1/9 & 2/9 \\ 1/9 & -2/9 & 2/9 \end{bmatrix} \tag{5}$$

Solution 3

Let A satisfy the characteristic equation $|A-\lambda I| = 0$. Since, $|A^T - \lambda I| = 0 \implies |(A-\lambda I)^T| = 0$, A and A^T have the same eigenvalues. So, we have to show that c is an eigenvalue of A.

$$A-\lambda I = \begin{bmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ & \dots & & \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{bmatrix} = \begin{bmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ & \dots & & \\ c-\lambda & c-\lambda & \dots & c-\lambda \end{bmatrix} \; (R_n \leftarrow R_1 + \dots + R_n)$$

$$|A-\lambda I| = 0 \implies |c-\lambda| \begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ & \dots & & \\ 1 & 1 & \dots & 1 \end{vmatrix} = 0 \tag{6}$$

Since, $\lambda = c$ is a solution of the characteristic equation, c is an eigenvalue of A, and hence A^T.

Solution 4

As per the Cayley–Hamilton theorem, every square matrix satisfies its characteristic equation.

$$|C-\lambda I| \implies \begin{vmatrix} c_{11}-\lambda & c_{12} \\ c_{21} & c_{22}-\lambda \end{vmatrix} = 0 \implies \lambda^2 - \lambda(c_{11}+c_{22}) + (c_{11}c_{22} - c_{12}c_{21}) = 0$$

Put $\lambda = C \implies C^2 - C \times trace(C) + |C| \; I_{2\times 2} = 0 \implies C^2 = -|C| \; I_{2\times 2}$ (given $trace(C) = 0$)

$$\implies C^n = \begin{cases} (-1)^{\frac{n-1}{2}} \; det(C)^{\frac{n-1}{2}} \; C & , n \text{ is odd} \\ (-1)^{\frac{n}{2}} \; det(C)^{\frac{n}{2}} \; I_{2\times 2} & , n \text{ is even} \end{cases} \tag{7}$$

Solution 5

(a) In general, the equality will not hold. An example is given as follows. Let us select two matrices C and D such that $C^n = D^n = 0 \; \forall \; n \geq 2$.

$$C = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad , \quad D = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \quad , \quad C + D = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

$$e^C = \sum_{n=0}^{\infty} \frac{C^n}{n!} = I + C = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad e^D = I + D = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \implies e^C e^D = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} \tag{8}$$

$$\text{Now, } (C+D)^{2n} = (-1)^n I, \; (C+D)^{2n+1} = (-1)^n (C+D)$$

$$e^{(C+D)} = \sum_{n=0}^{\infty} \frac{(C+D)^{2n}}{2n!} + \sum_{n=0}^{\infty} \frac{(C+D)^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n!} \; I + \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \; (C+D)$$

$$\implies e^{(C+D)} = cos(1)\ I + sin(1)\ (C+D) = \begin{bmatrix} cos(1) & -sin(1) \\ sin(1) & cos(1) \end{bmatrix} \tag{9}$$

Hence, from the above results, $e^{(C+D)} \neq e^C e^D$.

(b) In general, the product of two matrices C and D is not commutative, i.e. $CD \neq DC$. However, if one of the matrices is a multiple of the identity matrix, then $CD = DC$ as $CI = IC$. Hence, in this case, $e^{(C+D)} = e^C e^D$ condition will hold true.

Solution 6

Given, that $\vec{u}, \vec{v}$ are two vectors in $\mathbb{R}^2$ representing two sides of a triangle. Let the third side be $\vec{u} - \vec{v}$. So, we can obtain the following result using the cosine rule.

$$||\vec{u} - \vec{v}||^2 = ||\vec{u}||^2 + ||\vec{v}||^2 - 2\ ||\vec{u}||\ ||\vec{v}||\ cos\theta \tag{10}$$

$$\text{Also, } ||\vec{u} - \vec{v}||^2 = (\vec{u} - \vec{v}).(\vec{u} - \vec{v}) = ||\vec{u}||^2 + ||\vec{v}||^2 - 2\vec{u}.\vec{v} = ||\vec{u}||^2 + ||\vec{v}||^2 - 2 < \vec{u}, \vec{v} > \tag{11}$$

$$\implies ||\vec{u}||^2 + ||\vec{v}||^2 - 2\ ||\vec{u}||\ ||\vec{v}||\ cos\theta = ||\vec{u}||^2 + ||\vec{v}||^2 - 2 < \vec{u}, \vec{v} > \implies < \vec{u}, \vec{v} >= ||\vec{u}||\ ||\vec{v}||\ cos\theta$$

Solution 7 - Linear Dependence Lemma

If $U = (u_1, u_2, ..., u_n)$ are linearly dependent, then $c_1u_1 + c_2u_2 + ... + c_nu_n = 0$, where $c_1, c_2, ..., c_n$ are some constants. For $i \in \{2, 3, ..., n\}$ to be the largest index where $c_i \neq 0$,

$$u_i = -\frac{c_1}{c_i}u_1 - \frac{c_2}{c_i}u_2 - ... - -\frac{c_{i-1}}{c_i}u_{i-1}$$

Then u_i forms a vector space whose span is defined as $u_i \in span(u_1, u_2, ..., u_{i-1})$ and the basis $(u_1, u_2, ..., u_{i-1})$.

Now, consider $U' = U \setminus \{u_i\}$. Then, $span(U') \subset span(U) = s$ as $U' \subset U$.

$$s = d_1u_1 + d_{i-1}u_{i-1} + d_iu_i + ... + d_nu_n$$

Put the expression for u_i from above and we have that $s \in span(U') \implies span(U') \supset span(U)$. Since, the spans of both U and U' are each other's subsets, then we can conclude that their spans are equivalent, even after the removal of u_i.

Solution 8 - Independent Lists Cannot Be Arbitrarily Long

Let $U = u_1, u_2, ..., u_m$ be linearly independent in the vector space V. Also, $V = v_1, v_2, ..., v_n$ spans V. It needs to be shown that $m \leq n$. This can be done by adding one of u's and removing one of v's for multiple iterations.

Let us assume that $m > n$. Since B spans V, adding any vector in V to this list produces a linearly dependent list $v_1, v_2, ..., v_n, u_n$. Here, u_n can be obtained by a linear combination of vectors in B.

As per the Linear Dependence Lemma, if we remove one v_i, then the new list consisting of u_n and the remaining v's, will span V.

We will not be able to continue if there are no more u_i to add or v_i to remove (based on m and n). If $(m > n)$ then $span(v_1, v_2, ..., v_n) = V$ and $v_n \notin span(v_1, v_2, ..., v_n) = V$. But, it is contradictory to the initial statement that $v_i \in V \ \forall \ i$. Therefore, m must be $m \leq n$.

Hence, the length of any independent list of vectors is always smaller than or equal to the length of any spanning list of vectors.

Solution 9 - Cyclic Differences

---

1. Let B be linearly dependent. Then, the following should hold true.

$$c_1(u_1 - u_2) + c_2(u_2 - u_3) + ... + c_{n-1}(u_{n-1} - u_n) + c_n u_n = 0$$

$$\text{rearranging} \implies c_1 u_1 + (c_2 - c_1)u_2 + ... + (c_n - c_{n-1})u_n = 0$$

Now, linear independence of $(u_1, u_2, ..., u_n) \implies c_1 = c_2 - c_1 = ... = c_n - c_{n-1} = 0$. Hence, $c_i = 0 \ \forall \ i \in \{1, 2, ..., n\}$, so B is also linearly independent.

---

2. Let $u \in V$, therefore it can be written as a linear sum of the elements of the span of A.

$$c_1u_1 + c_2u_2 + ... + c_nu_n = u = d_1(u_1 - u_2) + d_2(u_2 - u_3) + ... + d_{n-1}(u_{n-1} - u_n) + d_n u_n \quad \text{(for B to span } V\text{)}$$

$$\text{rearranging RHS} \implies c_1u_1 + c_2u_2 + ... + c_nu_n = d_1u_1 + (d_2 - d_1)u_2 + ... + (d_n - d_{n-1})u_n$$

Hence, $c_1 = d_1, c_2 = d_2 - d_1, ..., c_n = d_n - d_{n-1} \implies d_k = \sum_{i=1}^{k} c_i$. Now, equation for B becomes,

$$B : c_1u_1 + (c_2 + c_3)(u_2) + ... + \left(\sum_{i=1}^{n} c_i \right) u_n = u \in span(A) = V \implies B \ spans \ V.$$

Solution 10

Let $A = \begin{bmatrix} 4 & -2 & 1 \\ 2 & 3 & -6 \\ 1 & 8 & 9 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & -2 & 1 \\ 8 & 3 & -6 \\ 9 & 2 & 1 \end{bmatrix}$. The $L1$ and *Frobenius* norms are as follows.

$$||A||_1 = max((|4| + |2| + |1|), (|-2| + |3| + |8|), (|1| + |-6| + |9|)) = max(7, 13, 16) = 16$$

$$\begin{aligned}
||B||_1 &= max((|4| + |8| + |9|), (|-2| + |3| + |2|), (|1| + |-6| + |1|)) = max(21, 7, 8) = 21 \\
||A||_F &= \sqrt{(|4|^2 + |2|^2 + |1|^2) + (|-2|^2 + |3|^2 + |8|^2) + (|1|^2 + |-6|^2 + |9|^2))} \\
&= \sqrt{16 + 4 + 1 + 4 + 9 + 64 + 1 + 36 + 81} = \sqrt{216} \\
||B||_F &= \sqrt{(|4|^2 + |8|^2 + |9|^2) + (|-2|^2 + |3|^2 + |2|^2) + (|1|^2 + |-6|^2 + |1|^2))} \\
&= \sqrt{16 + 64 + 81 + 4 + 9 + 4 + 1 + 36 + 1} = \sqrt{216}
\end{aligned}$$

Hence, $||A||_1 < ||B||_1$ and $||A||_F = ||B||_F$.

Solution 11 - Induced Matrix Norms

(a) We know that $||Ax|| \geq 0 \; \forall \; x$. If $||x|| \geq 0$, then $||A|| \; ||x|| \geq 0 \implies ||A|| \geq 0$. Hence, proved.

(b) We introduce a vector x to deal with the real number α. Then the LHS, $||\alpha A|| = ||(\alpha A)x|| = ||\alpha(Ax)|| \leq |\alpha| \; ||Ax|| \leq |\alpha| \; ||A||$. Hence, proved.

(c) $||A + B|| = \max\limits_{||x||=1} ||(A + B)x|| \leq \max\limits_{||x||=1} (||Ax|| + ||Bx||) \leq \max\limits_{||x||=1} (||A|| \; ||x|| + ||B|| \; ||x||) \leq ||A|| + ||B||$. Hence, proved.

(d) We know that $||A||$ is calculated from max($||A||$), expressed over a unit sphere. So, $||Ax|| = 0 \iff A = 0$. Hence, proved.

(e) We add a vector x into the expression. $||AB|| \leq \max\limits_{x \neq 0} \dfrac{||ABx||}{||x||} \leq \max\limits_{x \neq 0} \dfrac{||A|| \; ||Bx||}{||x||} \leq \max\limits_{x \neq 0} \dfrac{||A|| \; ||B|| \; ||x||}{||x||} \leq ||A|| \; ||B||$. Hence, proved.

(f) We define the induced $L2$ norm of matrix A in terms of vector x and use the properties of Singular Value Decomposition (SVD) to substitute it.

$$||A||_2 = \sup_{x \neq 0} \frac{||Ax||_2}{||x||_2} = \sup_{x \neq 0} \frac{||U \sum V^T x||_2}{||x||_2} \quad \text{(use SVD to rewrite } A\text{)}$$

$$= \sup_{x \neq 0} \frac{||\sum V^T x||_2}{||x||_2} = \sup_{y \neq 0} \frac{||\sum y||_2}{||Vy||_2} \quad (U \text{ and } V \text{ are complex unitary, so unit norm})$$

$$= \sup_{y \neq 0} \frac{||\sum y||_2}{||y||_2} = \sup_{y \neq 0} \frac{(\sum_{i=1}^{n} \sigma_i^2 |y_i|^2)^{\frac{1}{2}}}{(\sum_{i=1}^{n} |y_i|^2)^{\frac{1}{2}}} = \sigma_{max}(A) \quad (\text{expand the norm, supremum will be } \sigma_{max})$$

Solution 12

Given, a real symmetric matrix ($S_{n\times n}$), satisfying $S = S^T$ (symmetry condition) and $Sx = \lambda x$ (characteristic equation). For verifying the linear independence of the eigen vectors (x_i), we need to analyse their orthogonality. Consider,

$$(Sx_1).x_2 = (\lambda_1 x_1).x_2 = \lambda_1(x_1.x_2)$$

Now, check the same expression in the matrix form of dot product.

$$(Sx_1).x_2 = (Sx_1)^T x_2 = x_1^T S^T x_2 = x_1^T S x_2 = x_1^T (Sx_2) = x_1^T(\lambda_2 x_2) = \lambda_2(x_1.x_2)$$

From the above evaluations, we got $\lambda_1(x_1.x_2) = \lambda_2(x_1.x_2)$. Since, eigenvalues are unique, $x_1.x_2 = 0$. To generalise, $x_i.x_j = 0 \ \forall \ i \neq j$. Let us see the dot product of a particular eigen vector x_i with a linear combination of all the eigen vectors.

$$x_i.(c_1x_1+c_2x_2+...+c_nx_n) = x_i.c_ix_i = c_i||x_i||^2 = 0 \implies c_i = 0 \quad (\text{linear independence and } ||x|| \neq 0)$$

We had chosen x_i arbitrarily. So, the above conclusion is true for all c_i, i.e. $c_1 = c_2 = ... = c_n = 0$. Therefore, the linear independence condition is satisfied. Hence, the eigen vectors of a real symmetric ($S_{n\times n}$) matrix are linearly independent and form an orthogonal basis for $\mathbb{R}^n$.

Solution 13

The characteristic equation of a square matrix is $Ax = \lambda x$. Here, x is the eigen vector corresponding to the eigenvalue λ. So, for a matrix with n eigenvalues, there are n eigen vectors. We are given that $||x|| = 1$.

$$\max_x\{x^T A x\} = \max_x\{x_1^T A x_1, x_2^T A x_2, ..., x_n^T A x_n\} = \max_x\{x_1^T \lambda_1 x_1, x_2^T \lambda_2 x_2, ..., x_n^T \lambda_n x_n\}$$

$$\max_x\{\lambda_1 x_1^T x_1, \lambda_2 x_2^T x_2, ..., \lambda_n x_n^T x_n\} = \max_x\{\lambda_1||x_1||^2, \lambda_2||x_2||^2, ..., \lambda_n||x_n||^2\}$$

$$\max_x\{\lambda_1 \times 1, \lambda_2 \times 1, ..., \lambda_n \times 1\} = \max_x\{\lambda_1, \lambda_2, ..., \lambda_n\} \tag{12}$$

Hence, the solution is given by the largest eigenvalue of A, when x is the eigen vector corresponding to largest eigen value.

Solution 14

The matrix A has the non-zero eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ and the linearly independent eigen vectors $x_1, x_2, ..., x_n$, since it is full rank. Let us define a matrix $V_{n\times n}$ whose elements are these eigen vectors, where each individual vector is represented as a column matrix, $V = \begin{bmatrix} x_1 & x_2 & ... & x_n \end{bmatrix}$. Also, define an $n \times n$ diagonal matrix $\sum$, whose diagonal elements are the eigenvalues of A.

$$AV = A\begin{bmatrix} x_1 & x_2 & ... & x_n \end{bmatrix} = \begin{bmatrix} Ax_1 & Ax_2 & ... & Ax_n \end{bmatrix} = \begin{bmatrix} \lambda_1 x_1 & \lambda_2 x_2 & ... & \lambda_n x_n \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 x_{11} & \lambda_2 x_{21} & ... & \lambda_n x_{n1} \\ \lambda_1 x_{12} & \lambda_2 x_{22} & ... & \lambda_n x_{n2} \\ & ... & & \\ \lambda_1 x_{1n} & \lambda_2 x_{2n} & ... & \lambda_n x_{nn} \end{bmatrix} = \begin{bmatrix} x_{11} & x_{21} & ... & x_{n1} \\ x_{12} & x_{22} & ... & x_{n2} \\ & ... & & \\ x_{1n} & x_{2n} & ... & x_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ & ... & & \\ 0 & 0 & ... & \lambda_n \end{bmatrix} = V\sum$$

$$\implies AV = V\sum \implies AVV^{-1} = V\sum V^{-1} \implies A = V\sum V^{-1} \tag{13}$$

Hence, we can represent the full-rank matrix A in terms of the diagonal matrix $\sum$ and the matrix V formed by the eigen vectors.

Solution 15

Before transformation, $x.y = ||x||\ ||y||\ cos\theta = x^T y$ (in matrix form). Given that x is transformed into Ax and y is transformed into Ay, with new angle θ'. So, their inner product is given by,

$$(Ax).(Ay) = (Ax)^T(Ay) = x^T A^T A y$$

Since, A is an orthonormal transformation, $A^T A = I_{n\times n}$.

$$\implies (Ax).(Ay) = ||x||\ ||y||\ cos\theta'$$

Comparing the dot product before and after orthonormal transformation hence, $cos(\theta') = cos(\theta)$.

Solution 16

(a) We make use of two theorems namely Rank-nullity and Nullity theorem. Let $A = u_1 v_1^T$. Since, the column space of u_1 is contained in A, $rank(A) \leq rank(u_1)$.

Also, the null space of v_1^T is contained in the null space of A, as per the Rank-nullity theorem. Therefore, $rank(A) \leq rank(v_1^T)$ and $dim(N(A)) \geq dim(N(v_1^T))$.

Considering the constraints on $rank(A)$ and $dim(N(A))$ as described above, we apply the Nullity theorem to obtain $rank(A) = rank(u_1) = rank(v_1^T)$. Since, u_1, v_1 are orthonormal vectors ($\therefore$ rank 1). Hence, $rank(A) = rank(u_1 v_1^T) = 1$.

(b) $A = u_1v_1^T + u_2v_2^T$ can be thought to have been simplified version of $A = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix}$. We can use Singular Value Decomposition (SVD) to represent A in terms of rectangular (U, V) and diagonal $(\sum)$ matrices.

$$A = U \sum V^T = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix}$$

Here, the diagonal matrix $\sum$ ($I_{2\times2}$ in this case) dictates the maximum rank of A. So, the maximum $rank(A) = rank(u_1v_1^T + u_2v_2^T) = 2$ (since I has the dimensions 2×2).

(c) Again, we can use a similar approach as before. Let $A = \sum_{i=1}^{n} u_iv_i^T$ and apply SVD.

$$A = U \sum V^T = \begin{bmatrix} u_1 & u_2 & ... & u_n \end{bmatrix} \begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ & ... & & \\ 0 & 0 & ... & 1 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ ... \\ v_n^T \end{bmatrix}$$

Here, the diagonal matrix $\sum$ ($I_{n\times n}$ in this case) dictates the maximum rank of A. So, the maximum $rank(A) = rank(\sum_{i=1}^{n} u_iv_i^T) = n$ (since I has the dimensions $n \times n$).

References

[1] Mathematics for Machine Learning | Companion webpage to the book.
https://mml-book.com/

[2] Linear Algebra | Mathematics | MIT OpenCourseWare.
https://ocw.mit.edu/courses/18-06-linear-algebra-spring-2010/

[3] Singular Value Decomposition | Wikipedia.
https://en.wikipedia.org/wiki/Singular_value_decomposition

[4] Matrix Norm | Wikipedia.
https://en.wikipedia.org/wiki/Matrix_norm

■ ■ ■