Math, Inner Product Space#2

OSM Lab instructor,Zachary Boyd

OSM Lab student, CHEN Anhua

Due Wednesday, July 5 at 8:00am

1. (3.1)
   Define $< x, y >$ as $x \times y$ in a real inner porduct space. It's easy to show that $< x, y >$ satisfies all there properties of an IPS.
   i. RHS $= \frac{1}{4}((x+y)\times(x+y)-(x-y)\times(x-y)) = x \times y =$LHS
   ii. RHS $= \frac{1}{2}(2x^2+2y^2) = x^2+y^2 =$ LHS

2. (3.2)
   Let $x = a + bi$ and $y = c + di$, where $a, b, c, d \in \mathbb{R}$.
   LHS $= ac - bd + (bc + ad)i$
   RHS $= \frac{1}{4} \times (4ac - 4bd + 4adi + 4bci) =$LHS

3. (3.3)
   i. $cos\theta = \frac{<f,g>}{||f||||g||} = \frac{\int_0^1 x^6 dx}{\int_0^1 x^2 dx \int_0^1 x^1 0dx}$
   $= \frac{\sqrt{33}}{7}$
   ii. $cos\theta = \frac{<f,g>}{||f||||g||} = \frac{\int_0^1 x^6 dx}{\int_0^1 x^4 dx \int_0^1 x^8 dx}$
   $= \frac{3\sqrt{5}}{7}$

4. (3.8)
   i. for $x \in \{cos(t), sin(t), cos(2t), sin(2t)\}$, we always have $< x, x >= 1$. For $x, y \in \{cos(t), sin(t), cos(2t), sin(2t)\}$ and $x \neq y$, we always have $< x, y >= 0$. Therefore, S is an orthonormal set.
   ii. $||t|| =< t, t >= \frac{1}{\pi}\int_{\pi}^{\pi} t^2 dt = \frac{1}{\pi}(\frac{1}{3}\pi^3 - \frac{1}{3}(-\pi)^3) = \frac{2\pi^2}{3}$
   iii. Since S is an orthonormal set, for $x_i \in \{cos(t), sin(t), cos(2t), sin(2t)\}, proj_X cos(3t) = \sum_{x_i \in S} < cos(3t), x_i > x_i =$
   iv. Since S is an orthonormal set, for $x_i \in \{cos(t), sin(t), cos(2t), sin(2t)\}, proj_X t = \sum_{x_i \in S} < t, x_i > x_i =$

5. (3.9)
   For $x, y \in \mathbb{R}, M[x, y]^T = \frac{1}{\sqrt{5}}[x + 2y, 2x - y]^T$. Therefore, for $x, y, z, w \in \mathbb{R}$, $< M[x, y]^T, M[w, z]^T >= (xw + yz)$. We notice that $< [x, y]^T, [w, z]^T >= (xw + yz)$. Therefore, the rotation in $\mathbb{R}^2$ is an orthonormal transformation.

6. (3.10)
i. "if": Let $x, y \in \mathbb{F}^n$, according to Definition 3.2.14, $< Qx, Qy >= (Qx)^H Qy = x^H Q^H Qx =< x, y >= x^H y$. So $Q^H Q = I$. Trivial to prove that $QQ^H = I$
"only if": $< Qx, Qy >= (Qx)^H (Qy) = x^H Q^H Qy = x^H y =< x, y >$
ii. $< Qx, Qx >=< x, x >$. Therefore, $||Qx|| = ||x||$ for all $x \in mathbbF^n$
iii. First, $< Q^H x, Q^H y >= x^H (QQ^H) y =< x, y >$, therefore, Q^H is an orthonormal matrix. Also, $QQ^H = I \implies Q^{-1} QQ^H = Q^{-1} I \implies Q^{-1} = Q^H$. Therefore, Q^{-1} is an orthonormal matrix.
iv. Let $Q = \begin{bmatrix} q_1 & \dots & q_n \end{bmatrix}$ then $Q^H Q = \begin{bmatrix} q_1^H q_1 & \dots & q_1^H q_n \\ \vdots & \ddots & \vdots \\ q_n^H q_1 & \dots & q_n^H q_n \end{bmatrix} = I$
Therefore, the columns are orthonormal.
v. $1 = det(I) = det(QQ^H) = det(Q) det(Q^H)$. Because $det(Q) = det(Q^H)$, $|det(Q)| = 1$

7. (3.11)
W.L.O.G., let's assume $x_t = \sum_{j=1}^{t-1} \alpha_j x_j$. Then we have:
$p_{t-1} = \sum_{i=1}^{t-1} < q_i, x_t > q_i$
$= \sum_{i=1}^{t-1} < q_i, \sum_{j=1}^{t-1} \alpha_j x_j > q_i$
$= \sum_{i=1}^{t-1} \sum_{j=1}^{t-1} (\alpha_j < q_i, x_j > q_i)$
$= \sum_{i=1}^{t-1} \alpha_i x_i = x_t$
This means q_t is divided by 0 and G-S algorithm cannot proceed further on.

8. (3.16)
i. Let $A \in \mathbb{M}_{m \times n} A = QR$. Let D be a diagonal matrix with -1 on its first diagonal element and $1s$ on other diagonal positions. Then $A = Q(DD^{-1})R = (QD)(D^{-1}R)$. We find that QD is also an orthonormal matrix and $D^{-1}R$ is also an upper triangular matrix. Therefore, the QR decomposition is not unique.
ii. Assume that $A = Q_1 R_1 = Q_2 R_2 \implies R_1 R_2^{-1} = Q_1^H Q_2$. If we let $M = R_1 R_2^{-1} = Q_1^H Q_2$. We find tha M is both orthonormal and upper triangular, therefore a diagonal matrix. If diagonal elements of M are all positive, then $M = I$. Decomposition is unique.

9. (3.17)
$x = (A^H A)^{-1} A^H b$
$x = (\hat{R}^H (\hat{Q}^H \hat{Q}) \hat{R})^{-1} A^H b$
$x = \hat{R}^{-1} (\hat{R}^H)^{-1} \hat{R}^H \hat{Q}^H b$
$x = \hat{R}^{-1} \hat{Q}^H b$

10. (3.23)
i. $||x|| = ||x - y + y|| \leq ||x - y|| + ||y||$. Therefore, $||x|| - ||y|| \leq ||x - y||$.

ii. $||y|| = ||y-x+x|| \leq ||y-x|| + ||x||$. Therefore, $||y|| - ||x|| \leq ||y-x|| = ||-1|| \cdot ||x-y|| = ||x-y||$.
Combining i and ii, we could prove the question.

11. (3.24)
i. Positivity and scalar preservation are easy to check, therefore omitted; For triangular inequality, $||f+g||_{L^1} = \int_a^b |f(t)+g(t)|dt \leq \int_a^b (|f(t)|+|g(t)|)dt = ||f||_{L^1} + ||g||_{L^1}$
ii. Positivity and scalar preservation are easy to check, therefore omitted; For triangular inequality, it's easy to prove using the Minkowski's integral inequality.
iii. Positivity and scalar preservation are easy to check, therefore omitted; For triangular inequality, $\sup_{x\in[a,b]} |f(x)+g(x)| \leq \sup_{x\in[a,b]} |f(x)| + \sup_{x\in[a,b]} |g(x)|$. Therefore $||f+g||_\infty \leq ||f||_\infty + ||g||_\infty$.

12. (3.26)
(Equivalence relation) $\exists 0 < m \leq M s.t. m||x||_a \leq ||x||_b \leq M||x||_a$. Also $\exists 0 < \frac{1}{M} \leq \frac{1}{m} s.t. \frac{1}{M}||x||_b \leq ||x||_a \leq \frac{1}{m}||x||_b$. Therefore, it's an equivalent relation.
i. $||x||_2 = (\sum_{i=1}^n |x_i|^2)^{\frac{1}{2}} \leq ((\sum_{i=1}^n |x_i|)(\sum_{i=1}^n x_i))^{\frac{1}{2}} = \sum_{i=1}^n |x_i| = ||x||_1$. Also, $||x||_1 = \sum_{i=1}^n |x_i| \cdot 1 \leq (\sum_{i=1}^n |x_i|^2)^{\frac{1}{2}} (\sum_{i=1}^n 1^2)^{\frac{1}{2}} = \sqrt{n}||x||_2$
ii. $||x||_\infty = \sup\{|x_1|, \ldots, |x_n|\}$. $||x||_\infty \leq ||x||_2$ is trivial to prove. $||x||_2 = (\sum_{i=1}^n |x_i|^2)^{\frac{1}{2}} \leq= (\frac{1}{n}||x||_\infty^2)^{\frac{1}{2}} = \frac{1}{\sqrt{n}}||x||_\infty$

13. (3.28)
Since $||A||_p = \sup_{x\neq 0} \frac{||Ax||_p}{||x||_p}$, We could directly use te results from 3.26 to prove this question.
i. Given i) in 3.26, we know that $||Ax||_2 \leq ||Ax||_1 \leq \frac{1}{\sqrt{n}}||Ax||_2$ and $||x||_2 \leq ||x||_1 \leq \frac{1}{\sqrt{n}}||x||_2$. Therefore, $\frac{1}{\sqrt{n}}||A||_2 \leq ||A||_1 \leq \sqrt{n}||A||_2$
ii. Given ii) in 3.26, using the same logic in i) above, it's easy to prove.

14. (3.30)
i. The positivity is easy to prove.
ii. Homogneity: $||aA||_S = ||S(aA)S^{-1}|| = a||SAS^{-1}|| = a||A||_S$
iii. Triangular inequality: $||A+B||_S = ||S(A+B)S^{-1}|| = ||SAS^{-1} + SBS^{-1}||$. Given $||\cdot||$ is a matrix norm, it satisfies the triangular inequality. Therefore $||SAS^{-1} + SBS^{-1}|| \leq ||SAS^{-1}|| + ||SBS^{-1}|| = ||A||_S + ||B||_S$
Therefore, $||\cdot||_S is a matrix norm.$

15. (3.37)
suppose $p = a + bx + cx^2 = [1, x, x^2][a, b, c]^T$. Since$L(p) = p^{'}(1) = 2c + b =<$

$q, p >$. Therefore, $q = [\frac{2c+b}{a}, 0, 0]$

16. (3.38)
By using the same method in 3.37, we have $D = [b, a2c, 0]^T$

17. (3.39)
i. $< (S+T)^*y, x >=< y, (S+T)x >=< y, Sx > + < y, Tx >=< S^*y, x > + < T^*y, x >=< (S^* + T^*)y, x >$. So $(S+T)^* = (S^* + T^*)$. Also $< (\alpha T)^*y, x >=< y, \alpha Tx >= \overline{\alpha} < T^*y, x >=< \overline{\alpha}T^*y, x >$
. Therefore, $(\alpha T)^* = \overline{\alpha}(T)^*$.
ii. $< (S^*)^*y, x >=< y, S^*x >=< Sy, x >$, therefore, $(S^*)^* = S$
iii. $< (ST)^*y, x >=< y, (ST)x >=< S^*y, Tx >=< T^*S^*y, x >$. Therefore, $(ST)^* = T^*S^*$.
iv. AAccording to (iii), $(T^{-1})^*T^* = (TT^{-1})^* = I^* = I$. Therefore, $(T^{-1})^* = (T^*)^{-1}$.

18. (3.40)
i. Let $X, Y \in M_n(\mathbb{F})$, then $< Y, A^*X >=< A^*Y, X > \implies Y^H A^* X = (AY)^H X = Y^H A^H X$. Therefore, $A^* = A^H$
ii. Based on the equivalence proved in i, we have $< A_2, A_3A_1 >= tr(A_2^H A_3 A_1) = tr(A_2^* A_3 A_1) = tr(A_1 A_2^* A_3) = tr(A_1 A_2^H A_3) =< A_2 A_1^H, A_3 >=< A_2 A_1^*, A_3 >$
iii. $< Y, T_A^*(X) >=< T_A(Y), X >=< AY - YA, X >=< AY, X > - < YA, X >= tr(Y^H A^H X) - tr(A^H Y^H X) = tr(Y^H(A^H X)) - tr(Y^H(XA^H)) =< Y, A^H X > - < Y, XA^H >=< Y, A^*X > - < Y, XA^* >=< Y, T_{A^*}(X) >$. Therefore, $(T_A)^* = T_{A^*}$.

19. (3.44)
i. If $Ax = b$ has a solution $x \in \mathbb{F}$ then $b \in \mathcal{R}(A)$. Since $y \in \mathcal{N}(A^H) = \mathcal{N}(A^*) = \mathcal{R}(A)^\perp$, this means $< y, b >= 0$.
ii. If $Ax = b$ doesn't have a solution $x \in \mathbb{F}$, then it indicates that $b \in \mathcal{R}(A)^\perp$. Then there exists $y \in \mathcal{N}(A^H)$ such that $< y, b >\neq 0$

20. (3.45)
Given Theorem 3.8.5, to prove $Sym_n(\mathbb{R})^\perp = Skew_n(\mathbb{R})$ is equivalent to prove that $Sym_n(\mathbb{R}) \bigoplus Skew_n(\mathbb{R}) = \mathbb{R}$. Since $\forall A \in M_n(\mathbb{R})$ we could write $A = \frac{A+A^T}{2} + \frac{A-A^T}{2}$. It's easy to prove that $\frac{A+A^T}{2} \in Sym_n(\mathbb{R})$ and $\frac{A-A^T}{2} \in Skew_n(\mathbb{R})$. Therefore this decomposition holds for all satisfied A in the space.

21. (3.46)
i. By definition, $Ax \in \mathcal{R}(A)$. Since $x \in \mathcal{N}(A^H A)$, $A^H Ax = 0$. So $Ax \in \mathcal{N}(A^H)$.
ii. It's easy to prove that $\mathcal{N}(A) \subset \mathcal{N}(A^H A)$. If $x \in \mathcal{N}(A^H A)$, then $x^H A^H Ax =$

$0 \implies (Ax)^T(Ax) = 0 \implies Ax = 0 \implies x \in \mathcal{N}(A)$. So $\mathcal{N}(A^HA) = \mathcal{N}(A)$.
iii. Combininig ii and Rank-Nullity Theorem, it's easy to prove, given $\mathcal{N}(A^HA) = \mathcal{N}(A)$
iv. Natural to prove directly from arguments in iii.

22. (3.47)
i. $P^2 = A(A^HA)^{-1}A^HA(A^HA)^{-1}A^H = A(A^HA)^{-1}A^H = P$
ii. $P^H = (A(A^HA)^{-1}A^H)^H = A(A^HA)^{-1}A^H = P$
iii. $Rank(P) = Rank(A(A^HA)^{-1}A^H) \leq Rank(A)$, and $Rank(A) = Rank(A(A^HA)^{-1}(A^HA)) = Rank(PA) \leq Rank(P)$. Therefore, $Rank(A) = Rank(P)$

23. (3.48)
i. $P(A+B) = \frac{(A+B)+(A+B)^T}{2} = \frac{(A+B)+A^T+B^T}{2} = P(A) + P(B)$
ii. $P^2(A) = P(\frac{1}{2}(A+A^T)) = \frac{\frac{1}{2}(A+A^T)+\frac{1}{2}(A+A^T)^T}{2} = P(A)$
iii. We could directly using the fact that $P^* = P^T$ in question i in 3.40 and$P^T = P$ to prove this.
iv. $\forall x \in \mathcal{N}(P), Px = 0 \iff \frac{x+x^T}{2} = 0 \iff x^T = -x \iff x \in Skew_n(\mathbb{R})$. Therefore, $\mathcal{N}(P) = Skew_n(\mathbb{R})$
v. By using the result of 3.45 and question (iii), we know that $Sym_n(\mathbb{R}) = Skew_n(\mathbb{R})^\perp = \mathcal{R}(P^*) = \mathcal{R}(P)$.
vi. $A - P(A) = \frac{A-A^T}{2}$. Therefore, $\sqrt{< \frac{A-A^T}{2}, \frac{A-A^T}{2} >} = \sqrt{\frac{1}{4}(tr(A^TA) - tr(A^2) - tr((A^T)^2)))} = \sqrt{\frac{tr(A^TA) - tr(A^2)}{2}}$

24. (3.50)
$$A = \begin{bmatrix} x_1^2 & 1 \\ x_2^2 & 1 \\ \vdots & \vdots \\ x_n^2 & 1 \end{bmatrix}$$
$$x = \begin{bmatrix} -\frac{r}{s} \\ \frac{1}{s} \end{bmatrix}$$
$$b = \begin{bmatrix} y_1^2 \\ \vdots \\ y_n^2 \end{bmatrix}$$