MATB24 TUTORIAL PROBLEMS 7, SOLUTIONS

KEY WORDS: inner product, inner product space, dot product, orthogonal complement, orthogonal projection

RELEVANT SECTIONS IN THE TEXTBOOK: Sec 3.5, 6.1, 6.2

WARM-UP:

(1) Write down a complete definition or a complete mathematical characterization for the following terms.
- An inner product on a vector apace V
- An inner product space
- Orthogonal complement of a subspace W of an inner product space V
- Orthogonal decomposition of a vector $\vec{v}$ in an inner product space V
- An orthogonal set of vectors
- An orthogonal basis for a vector space W
- An orthonormal basis for a vector space W

(2) Give an example of the described object or explain why such an example does not exists.
- An inner product on $\mathbb{R}^2$ other than the dot product.
- An inner product on $\mathbb{R}^3$ other than the dot product.
- Two vectors in $\mathbb{R}^3$ that are orthogonal with respect to dot product but not with respect to your example of inner product.
- A vector in $\mathbb{R}^3$ with length one with respect to dot product and a different length with respect to your example of inner product.
- Two different orthogonal bases of $\mathbb{R}^2$.
- A vector in an inner product space V that is orthogonal to every other vector.
- 4 mutually orthogonal vectors in $\mathbb{R}^3$.

A:

(1) Let V_1 be the subspace of $\mathbb{R}^4$ given by

$$\begin{aligned} x_1 - x_2 - 2x_3 &= 0, \\ x_2 + x_3 - 2x_4 &= 0. \end{aligned}$$

Find an orthonormal basis of V_1.

(2) Let V_2 be the subspace of $\mathbb{R}^4$ given by

$$x_1 + x_2 - x_3 - 2x_4 = 0.$$

Find an orthonormal basis of V_2.

(3) Let $\vec{w}$ be the vector

$$\vec{w} = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 2 \end{bmatrix} \in \mathbb{R}^4.$$

Find the orthogonal projections of $\vec{w}$ onto the subspaces V_1, V_2.

(4) Show that any subspace V of $\mathbb{R}^n$ has an orthonormal basis. [1]

[1] Hint: start with any basis, then try to modify it to make it orthonormal.

Solution.

(1) Writing the equations as rows yields the matrix $\begin{bmatrix} 1 & -1 & -2 & 0 \\ 0 & 1 & 1 & -2 \end{bmatrix}$. Letting the last two columns be the free variables r, s, the nullspace is given by $[(-s+2r)+2s, -s+2r, s, r]$, which is spanned by the following two vectors:

$$(\vec{u}, \vec{v}) \ = \ \left(\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 0 \\ 1 \end{bmatrix} \right),$$

These are already orthogonal, so normalizing gives the orthonormal basis $\left(\frac{1}{\sqrt{3}}\vec{u}, \frac{1}{3}\vec{v}\right)$.

(2) By inspection, $\vec{u} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}$ are orthogonal vectors in V_2, and $\vec{z} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix} \in V_2$ is independent of $\vec{u}$ and $\vec{v}$ (e.g. the zeros in components do not match) and orthogonal to $\vec{u}$. So

$$(\vec{u}, \vec{v}, \vec{z} - \text{proj}_{\vec{v}}(\vec{z})) \ = \ \left(\vec{u}, \vec{v}, \vec{z} - \frac{\vec{z}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\vec{v}\right) \ = \ \left(\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -\frac{2}{5} \\ -1 \\ \frac{4}{5} \end{bmatrix} \right)$$

is an orthogonal set in V_2 (e.g. $\vec{z}_{\vec{v}^\perp} = z - \text{proj}_{\vec{v}}(\vec{z})$, so this vector is orthogonal to $\vec{v}$, and since $\text{proj}_{\vec{v}}(\vec{z})$ is parallel to $\vec{v}$ which is orthogonal to $\vec{u}$, and since $\vec{z}$ is orthogonal to $\vec{u}$, therefore $\vec{z}_{\vec{v}^\perp} \cdot \vec{u} = \vec{z}\cdot\vec{u} - \text{proj}_{\vec{v}}(\vec{z}) \cdot u = 0 - 0 = 0$), and by normalizing we get an orthonormal basis of V_2.

(3) Letting $\vec{u}, \vec{v}$ be as in the solution to Problem 1, we have

$$\text{proj}_{V_1}(\vec{w}) \ = \ \left(\vec{w}\cdot\frac{1}{\sqrt{3}}\vec{u}\right)\frac{1}{\sqrt{3}}\vec{u} + \left(\vec{w}\cdot\frac{1}{3}\vec{v}\right)\frac{1}{3}\vec{v} \ = \ \frac{1}{9}\begin{bmatrix} 10 \\ 22 \\ -6 \\ 8 \end{bmatrix}.$$

By a similar computation using Problem (2),

$$\text{proj}_{V_2}(\vec{w}) \ = \ \begin{bmatrix} 1 \\ 2 \\ -1 \\ 2 \end{bmatrix}.$$

(4) This is the Gram-Schmidt process. Convert the basis $(\vec{b}_1, \dots, \vec{b}_r)$ of V into the orthonormal basis $(\vec{u}_1, \dots, \vec{u}_r)$ of V by inductively defining $\vec{u}_k$ to be the normalization of

$$\vec{b}_k - \sum_{i=1}^{k-1} \frac{\vec{b}_k\cdot\vec{u}_i}{\vec{u}_i\cdot\vec{u}_i}\vec{u}_i.$$

B: Let V be a subspace of $\mathbb{R}^m$. We proved in class that for every $\vec{x} \in \mathbb{R}^m$ there is a unique vector $\text{proj}_V(\vec{x}) \in V$ such that $\vec{x} - \text{proj}_V(\vec{x})$ is orthogonal to V. The vector $\text{proj}_V(\vec{x})$ is called the *orthogonal projection* of $\vec{x}$ onto V, and we found a formula for it: if $\mathcal{U} = \{\vec{u}_1, \ldots, \vec{u}_n\}$ is an orthonormal basis of V, then

$$\text{proj}_V(\vec{x}) = \sum_{i=1}^{n} (\vec{x} \cdot \vec{u}_i)\vec{u}_i.$$

(1) Prove that the transformation $\text{proj}_V : \mathbb{R}^m \to \mathbb{R}^m$ is linear.
(2) What is the kernel of proj_V?
(3) Write the rank nullity equation for this transformation. Is it familiar?
(4) Since the orthogonal projection $\text{proj}_V : \mathbb{R}^m \to \mathbb{R}^m$ onto V is linear, there is an $m \times m$ matrix P such that $\text{proj}_V(\vec{x}) = P\vec{x}$ for all $\vec{x} \in \mathbb{R}^m$. Can you find that matrix.[2]

Solution.

(1) Using the formula $\text{proj}_V(\vec{x}) = \sum_{i=1}^{n} (\vec{x} \cdot \vec{u}_i)\vec{u}_i$, this follows from the linearity of the dot product.
(2) $V^\perp$. For example, if $\text{proj}_V(\vec{x}) = \vec{0}$, then $\vec{x} = \vec{x} - \text{proj}(\vec{x})$ is orthogonal to V, i.e. $\vec{x} \in V^\perp$.
(3) Since $V = \text{Im}(\text{proj}_V)$ and $V^\perp = \ker(\text{proj}_V)$, the rank-nullity theorem says $\dim(V) + \dim(V^\perp) = m$
(4) If A is the $m \times n$ matrix with columns $\vec{u}_1, \ldots, \vec{u}_n$, then $\text{proj}(\vec{x}) = A(A^TA)^{-1}A^T\vec{x}$. The following is a brief explanation. Since $\text{proj}(\vec{x}) \in V$, and the column space of A is precisely V, therefore $\text{proj}(\vec{x}) = A\vec{r}$ for some $\vec{r} \in \mathbb{R}^n$. Then, since $\vec{x} - A\vec{r}$ is orthogonal to V, therefore for any $\vec{y} \in \mathbb{R}^n$, $(\vec{x} - A\vec{r}) \cdot A\vec{y} = 0$, i.e. $(A\vec{y})^T(\vec{x} - A\vec{r}) = \vec{y}^T(A^T\vec{x} - A^TA\vec{r}) = 0$. Since this holds for all such $\vec{y}$, it follows that $A^T\vec{x} - A^TA\vec{r} = 0$. Since A is invertible due to its columns being independent, therefore A^TA is invertible, and hence $\vec{r} = (A^TA)^{-1}A^T\vec{x}$. Then, $\text{proj}_V(\vec{x}) = A\vec{r} = A(A^TA)^{-1}A^T\vec{x}$.

C: **Theorem 1:** *Let A be an $m \times n$ matrix. Then for all $\vec{x} \in \mathbb{R}^n$ and $\vec{y} \in \mathbb{R}^m$, we have an equality*

$$A\vec{x} \cdot \vec{y} = \vec{x} \cdot A^T\vec{y}.$$

Theorem 2: If A is an $m \times n$ matrix, then $\ker A^T = (\text{Im}A)^\perp$. Also, $(\ker A^T)^\perp = (\text{Im}A)$.

(1) Verify Theorem 1 for the matrix $A = I_n$ and any $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^n$
(2) Verify Theorem 1 for the matrix $A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$ and any $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^2$.
(3) Prove Theorem 1[3]
(4) Prove theorem 2 using theorem 1.

Solution.

(1) $I_n\vec{x} \cdot \vec{y} = \vec{x} \cdot \vec{y} = \vec{x} \cdot I_n\vec{y}$ since $I_n^T = I_n$.
(2) $A\vec{x} \cdot \vec{y} = [x_1 + 2x_2, -x_2] \cdot y = x_1y_1 + 2x_2y_1 - x_2y_2 = \vec{x} \cdot [y_1, 2y_1 - y_2]$.

[2] we will do this in class

[3] by interpreting the dot product of two vectors $\vec{w}$ and $\vec{v}$ in $\mathbb{R}^n$ as a matrix product $\vec{w}^T\vec{v}$.

(3) $\vec{A}x \cdot \vec{y} = (A\vec{x})^T\vec{y} = (\vec{x}^T)A^T\vec{y} = \vec{x}^T(A^T\vec{y}) = \vec{x} \cdot A^T\vec{y}$.
(4) Take $x \in \ker A^T$. We need $y \cdot x = 0$ for all $y \in \mathrm{Im}A$. We need $Az \cdot x = 0$ for all z. By Theorem 1, this is the same as $z \cdot A^T x = 0$ for all z. Of course, this true because $A^T x = 0$ (def of kernel). For the reverse inclusion: say $y \in (\mathrm{Im}\, A)^{\perp}$. This means $Ax \cdot y = 0$ for all x. This is the same as $x \cdot A^T y = 0$ for all x. This means $A^T y = 0$, so $y \in kerA^T$. The second statement holds by "perping" both sides.

D:

Let V and W be subspaces of coordinate vector spaces $\mathbb{R}^n$ and $\mathbb{R}^m$, and suppose that $T : V \to W$ is an isomorphism.

(1) If $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_r\}$ is a basis of V, is the set $\{T(\vec{b}_1), \ldots, T(\vec{b}_r)\}$ necessarily a basis of W?
(2) If $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_r\}$ is an orthonormal basis of V, is the set $\{T(\vec{b}_1), \ldots, T(\vec{b}_r)\}$ necessarily an orthonormal basis of W?
(3) Which $n \times n$ matrices A have the property that for every orthonormal basis $\{\vec{u}_1, \ldots, \vec{u}_n\}$ of $\mathbb{R}^n$, the set $\{A\vec{u}_1, \ldots, A\vec{u}_n\}$ is also an orthonormal basis of $\mathbb{R}^n$? Can you come up with a few examples that works and a few that doesn't?

Solution.

(1) Yes!
(2) No! For instance, dilating $\mathbb{R}^2$ by a factor of 2 (i.e. $T(\vec{x}) = 2\vec{x}$) converts the orthonormal basis $(\vec{e}_1, \vec{e}_2)$ of $\mathbb{R}^2$ into the non-orthonormal basis $(2\vec{e}_1, 2\vec{e}_2)$.
(3) By taking $(\vec{u}_1, \ldots, \vec{u}_n)$ to be the standard basis of $\mathbb{R}^n$, we see that the columns of A must form an orthonormal basis of $\mathbb{R}^n$. In fact, this is also a sufficient condition, since a product of orthogonal $n \times n$ matrices is itself an orthogonal matrix. We will be better equipped to understand this after defining *orthogonal* transformations and matrices on the next week.

Exercises from the book: 1,3,5,7,9,11,19,20,24,25,26-28,33