Recitation 9

Thurs Sep 26, 2024

1 Recap

1.1 Orthogonality of Vectors

Let u and v be vectors in $\mathbb{R}^m$. Their dot ("inner") product "$u \cdot v$" can be written $u^T v = v^T u$ in linear-algebra notation (some people also use the notation $\langle u, v \rangle$). A matrix transpose (swapping rows and columns) has the key property that it *moves the matrix to the other side of a dot product*: $u \cdot (Av) = u^T Av = (A^T u)^T v = (A^T u) \cdot v$ for any u, v. Indeed, this can be treated as the *definition* of a matrix transpose (with swapping rows and columns being the *consequence*), and the practical applications of matrix transposes always involves a dot product somewhere.

We say that u and v are *orthogonal* if and only if their dot product is 0. Furthermore, we define a set of vectors $\{v_1, v_2, ..., v_n\}$ as *pairwise orthogonal* if and only if every pair of distinct vectors is orthogonal. In other words, v_i and v_j are orthogonal for any $i, j \in \{1, 2, ..., n\}$ such that $i \neq j$. A set of pairwise orthogonal nonzero vectors is always linearly independent.

A set of vectors $\{q_1, q_2, ..., q_n\}$ in $\mathbb{R}^n$ is *orthonormal* if it is pairwise orthogonal, and each q_i is a unit vector $\|q_i\| = 1$. ("q_k" in linear algebra almost always means that the vectors are orthonormal.).

If we put such orthonormal vectors into the columns of a matrix, we get the $m \times n$ matrix $Q = \begin{pmatrix} q_1 & q_2 & \cdots & q_n \end{pmatrix}$, and the orthonormality is *equivalent* to the property that $Q^T Q = I$ (the $n \times n$ identity). It does *not* follow that $Q^T = Q^{-1}$ unless Q is *square* (which we call an "orthogonal" or "unitary" matrix); for a non-square Q, $QQ^T \neq I$.

Two key factors that make orthonormal bases nice is that (1) finding their coefficients only involves dot products and (2) they preserve lengths. That is, if $x = Qc$, then the basis coefficients c are simply $c = Q^T x$, and also $\|x\| = \|c\|$.

1.2 Orthogonality of Subspaces

Two subspaces S_1 and S_2 of V are *orthogonal* if every vector in S_1 is orthogonal to every vector in S_2, meaning that $x^T y = 0$ for all $x \in S_1$ and $y \in S_2$. Furthermore, it follows that $\dim S_1 + \dim S_2 \leq \dim V$.

1.3 Orthogonal Complement of Subspaces

Given a subspace S of V, its *orthogonal complement* $S^\perp$ is defined as the subspace containing *all* vectors in V that are orthogonal to *every* vector in S. In other words,

$$S^\perp = \left\{ w \in V : w^T x = 0 \text{ for any } x \in S \right\}.$$

Therefore, $S^\perp$ is the largest subspace that is orthogonal to S. Here are some properties of the orthogonal complement:

1. $\dim S + \dim S^\perp = \dim V$
2. $\left(S^\perp\right)^\perp = S$

1.4 Some Familiar Orthogonal Complements

We have already seen and worked on orthogonal complements, but we just didn't realize that they are!

Theorem 1 *$N(A)$ and $C(A^T)$ are orthogonal complements in $\mathbb{R}^n$. Similarly, $C(A)$ and $N(A^T)$ are orthogonal complements in $\mathbb{R}^m$.*

2 Exercises

1. $A^T B \neq B^T A$ for general matrices (with compatible sizes), but $B^T A$ *is* the __________ of $A^T B$. However, for *column vectors*, $x^T y = y^T x$ for any $x, y \in \mathbb{R}^m$. Why does it work for any vectors but not any matrices?

2. If S is any subspace of V, then the *intersection* of S with $S^\perp$ must be __________?

3. We've already seen permutation matrices P: $m \times m$ matrices that re-order vectors in $\mathbb{R}^m$. Explain why any such P must have orthonormal columns (and rows), so it is an "orthogonal (unitary) matrix:" $P^{-1} = P^T$ (as we claimed without proof a few lectures ago!).

4. Denote a subspace $V = \left\{ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} : 2v_1 + 3v_2 + 5v_3 = 0 \right\}$. Find $V^\perp$ (give a basis). Can you relate V and $V^\perp$ to column and/or null spaces of some matrix and/or its transpose?

5. Suppose that we have a subspace $\mathcal{S}$ with an orthogonal (but *not* necessarily orthonormal) basis $\{v_1, ..., v_k\}$.

   (a) By the definition of basis, any vector $v \in \mathcal{S}$ can be expressed as

   $$v = \sum_{i=1}^{k} \alpha_i v_i = \alpha_1 v_1 + ... \alpha_k v_k$$

   for some constants $\alpha_1, ..., \alpha_k$. Determine each α_j in terms of $v_1, ..., v_k$ and v. How does it differ from the orthonormal case? Will the same derivation work without the orthogonality of $\{v_1, ..., v_k\}$?

   (b) If we put the vectors v_i as the columns of a matrix $V = \begin{pmatrix} v_1 & \cdots & v_k \end{pmatrix}$, what special form does $V^T V$ have?

3 Solutions

1. They are *transposes*: $B^T A = (A^T B)^T$ (as claimed in class; you will derive this in pset 3). So they are not equal unless $A^T B$ is symmetric (equal to its transpose). However, for vectors, $x^T y$ is a scalar (or "1×1" matrix), which *is* equal to its transpose, so $x^T y = y^T x$. You can also see this by writing out the dot product $x^T y = x_1 y_1 + \cdots + x_m y_m$, which clearly equals $y^T x$ since the elementwise multiplications are commutative. You are probably also familiar with the statement that $x \cdot y = y \cdot x$ in high-school vector algebra, first-year physics, and so on. But it is good to re-examine this—we never want to take commutativity for granted in linear algebra!

2. The intersection must be $\{\vec{0}\}$: the zero-dimensional subspace consisting only of the origin $\vec{0}$ of V. The reason for this is that $\vec{0}$ is the only vector that can be orthogonal to *itself*: $\vec{0}^T \vec{0} = \|\vec{0}\|^2 = 0$, which is not true for any other vector $x \neq \vec{0}$ since $x^T x = \|x\|^2 > 0$ for non-zero vectors. And being orthogonal to itself is the only way it can be in both S and $S^\perp$.

3. There are several ways to see this, for example:

   - The most straightforward approach: $PI = P$, but this means that P consists of *permuting the rows* of the identity matrix I. The rows of I are orthonormal, and re-ordering them doesn't change this, so P has orthonormal rows, hence $PP^T = I$. Since P is square this also means $P^T P = I$ (i.e. P has orthonormal columns, simply a permutation P^T of the columns of I). Hence $P^T = P^{-1}$: it is an orthogonal (unitary) matrix.
   - Here is a more subtle proof that uses the relationship of unitary matrices to inner products: Take any vectors $x, y \in \mathbb{R}^m$. Then we must have $x^T y = (Px)^T (Py)$, since re-ordering the rows of *both* x and y in the same way cannot change their dot product. That is, P "preserves" inner products. But $(Px)^T (Py) = x^T P^T P y$ for all x, y, which can only be true if $P^T P = I$.

4. There are several ways to solve this problem. First, we realize that V resides in an ambient space $\mathbb{R}^3$ and has dimension 2. This means $V^\perp$ has dimension $3 - 2 = 1$, which means that if we can find a non-zero vector $w \in V^\perp$, we can write $V^\perp = \text{Span}\{w\}$.

   Any $\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in V$ must satisfy $2v_1 + 3v_2 + 5v_3 = 0$, i.e. $\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = 0$. By definition, $V^\perp = \{w : w \cdot v = 0 \text{ for any } v \in V\}$. This implies that $w = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$ works, which means $V^\perp = \text{Span}\left\{ \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} \right\}$.

   Another way to solve this problem is to notice that $V = N(A)$ when $A = \begin{bmatrix} 2 & 3 & 5 \end{bmatrix}$. This means $V^\perp = C(A^T) = \text{Span}\left\{ \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} \right\}$.

5. (a) To derive α_j, we take inner product, for both sides of the equation, with v_j.

$$v_j^T v = \sum_{i=1}^{k} v_j^T(\alpha_i v_i) = \sum_{i=1}^{k} \alpha_i \underbrace{v_j^T v_i}_{0 \text{ if } i \neq j} = \alpha_j v_j^T v_j = \alpha_j \|v_j\|^2.$$

This implies $\alpha_j = \frac{v_j^T v}{\|v_j\|^2}$. We make a crucial note that this derivation only works when $\{v_1, ..., v_k\}$ are orthogonal, which we used to cancel the $i \neq j$ terms. In the orthonormal case, things simplify further because $\|v_j\|^2 = 1$.

(b) The i, j entry of $V^T V$ is $v_i^T v_j$. This means that $V^T V$ is a *diagonal* matrix of the form:

$$V^T V = \begin{pmatrix} \|v_1\|^2 & & & \\ & \|v_2\|^2 & & \\ & & \ddots & \\ & & & \|v_k\|^2 \end{pmatrix}$$

(which simplifies to the identity matrix if the basis is orthonormal).