Lecture 2 - *Orthogonal Vectors and Matrices*

OBJECTIVE: The best algorithms of numerical linear algebra are somehow based on orthogonality.
We review the ingredients: orthogonal vectors and matrices.

◇ TRANSPOSE

Definition 1. *The* transpose $\mathbf{A^T}$ *of an* $m \times n$ *matrix* $\mathbf{A}$ *is an* $n \times m$ *matrix where the (*i*,*j*) entry of* $\mathbf{A^T}$ *is the (*j*,*i*) entry of* $\mathbf{A}$
$\leftrightarrow$ *interchange the rows with the columns.*

e.g., If $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$, then $\mathbf{A^T} = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$.

If $\mathbf{A} = \mathbf{A^T}$ (so $\mathbf{A}$ has to be square!) then $\mathbf{A}$ is said to be *symmetric*.

Note 1. *The text uses* $\mathbf{A}^*$ *to denote* $\mathbf{A^T}$ *because it allows for complex numbers.*

◇ INNER PRODUCT

Definition 2. *Let* $\mathbf{x}$,$\mathbf{y}$ $\in$ $\mathbb{R}^m$. *Then, the* inner product *of* $\mathbf{x}$ *and* $\mathbf{y}$ *is a* <u>*scalar*</u>

$$\mathbf{x}^{\mathbf{T}}\mathbf{y} = \sum_{i=1}^{m} x_i y_i.$$

The (Euclidean) length *of a vector* $\mathbf{x}$ *is written as* $\|\mathbf{x}\|$ *and can be defined as the square root of the inner product of the vector with itself*

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^{\mathbf{T}}\mathbf{x}} = \left(\sum_{i=1}^{m} x_i^2\right)^{\frac{1}{2}}.$$

Also, if the angle between vectors $\mathbf{x}$ *and* $\mathbf{y}$ *is* α, *we have*

$$\cos\ \alpha = \frac{\mathbf{x}^{\mathbf{T}}\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}.$$

Some useful relationships:

a) $(\mathbf{x}_1 + \mathbf{x}_2)^T \mathbf{y} = \mathbf{x}_1^T \mathbf{y} + \mathbf{x}_2^T \mathbf{y}$

b) $\mathbf{x}^\mathbf{T}(\mathbf{y}_1 + \mathbf{y}_2) = \mathbf{x}^\mathbf{T}\mathbf{y}_1 + \mathbf{x}^\mathbf{T}\mathbf{y}_2$

c) $(\alpha\mathbf{x})^T(\beta\mathbf{y}) = \alpha\beta\mathbf{x}^\mathbf{T}\mathbf{y}$

Which is faster to compute, a) or b)?

The following properties are also true *provided the operations are defined!*

d) $(\mathbf{AB})^T = \mathbf{B}^\mathbf{T}\mathbf{A}^\mathbf{T}$

e) $(\mathbf{AB})^{-1} = \mathbf{B}^{-\mathbf{1}}\mathbf{A}^{-\mathbf{1}}$

Note 2. $\mathbf{A}^{-\mathbf{T}}$ *is shorthand for* $(\mathbf{A}^\mathbf{T})^{-1}$ *or* $(\mathbf{A}^{-\mathbf{1}})^T$*.* They are equal!

$\diamond$ ORTHOGONAL VECTORS

Vectors $\mathbf{x}, \mathbf{y}$ are orthogonal if $\mathbf{x}^{\mathbf{T}}\mathbf{y} = 0$;
i.e., the angle between $\mathbf{x}$ and $\mathbf{y}$ is $\frac{\pi}{2}$. (*verify!*)

Two sets of vectors $\mathbf{X}, \mathbf{Y}$ are (mutually) orthogonal if every $\mathbf{x} \in \mathbf{X}$ is orthogonal to every $\mathbf{y} \in \mathbf{Y}$.

One set of nonzero vectors $\mathbf{S}$ is (mutually) orthogonal if its elements are (pairwise) orthogonal
i.e., if $\mathbf{x}, \mathbf{y} \in \mathbf{S}$ and $\mathbf{x} \neq \mathbf{y}$, then $\mathbf{x}^{\mathbf{T}}\mathbf{y} = 0$.

A set of vectors $\mathbf{S}$ is *orthonormal* if it is orthogonal and for every $\mathbf{x} \in \mathbf{S}$, $\|\mathbf{x}\| = 1$.

Theorem 1. *The vectors in an orthogonal set* $\mathbf{S}$ *are linearly independent.*

Corollary 1. *If* $\mathbf{S} \subseteq \mathbb{R}^m$ *contains* m *vectors, then it is a basis for* $\mathbb{R}^m$.

$\diamond$ COMPONENTS OF A VECTOR

This is the most important idea:
Inner products can be used to decompose arbitrary vectors into orthogonal components.

e.g., Let $\mathbf{Q} = \{\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_n\}$ be an orthonormal set and let $\mathbf{v}$ be an arbitrary vector.

$\mathbf{q}_j^T \mathbf{v}$ is a scalar that represents the jth *coordinate* of $\mathbf{v}$ in basis $\mathbf{Q}$.

$\mathbf{v}$ can be decomposed into $n+1$ orthogonal components:

$$\mathbf{v} = \mathbf{r} + \sum_{i=1}^{n} (\mathbf{q}_i^T \mathbf{v}) \mathbf{q}_i = \mathbf{r} + \sum_{i=1}^{n} (\mathbf{q}_i \mathbf{q}_i^T) \mathbf{v}.$$

It is easy to verify that $\mathbf{r}$ is orthogonal to $\mathbf{Q}$:

$$\mathbf{r} = \mathbf{v} - (\mathbf{q}_1^T\mathbf{v})\mathbf{q}_1 - (\mathbf{q}_2^T\mathbf{v})\mathbf{q}_2 - \ldots - (\mathbf{q}_n^T\mathbf{v})\mathbf{q}_n$$

So,

$$\mathbf{q}_i^T\mathbf{r} = \mathbf{q}_1^T\mathbf{v} - (\mathbf{q}_1^T\mathbf{v})\mathbf{q}_i^T\mathbf{q}_1 - (\mathbf{q}_2^T\mathbf{v})\mathbf{q}_i^T\mathbf{q}_2 - \ldots - (\mathbf{q}_n^T\mathbf{v})\mathbf{q}_i^T\mathbf{q}_n$$

Since $\mathbf{q}_i^T\mathbf{q}_j = 0$ if $i \neq j$, the only term left is $(\mathbf{q}_i^T\mathbf{v})(\mathbf{q}_i^T\mathbf{q}_i)$.

$$\mathbf{q}_i^T\mathbf{r} = \mathbf{q}_i^T\mathbf{v} - (\mathbf{q}_i^T\mathbf{v})(\mathbf{q}_i^T\mathbf{q}_i) = 0$$

$\rightarrow$ $\mathbf{r}$ is the part of $\mathbf{v}$ orthogonal to $\mathbf{Q}$.

So, if $\mathbf{Q}$ is a basis for $\mathbb{R}^m$, $n = m$ and $\mathbf{r} = \mathbf{0}$. (why?)

$$\therefore \qquad \mathbf{v} = \sum_{i=1}^{m} (\mathbf{q}_i^T \mathbf{v}) \mathbf{q}_i = \sum_{i=1}^{m} (\mathbf{q}_i \mathbf{q}_i^T) \mathbf{v}$$

Note 3. *The sums are written in two different ways because there are two different interpretations.*

1. $(\mathbf{q}_i^T \mathbf{v})$ is the coordinate in direction $\mathbf{q}_i$

2. $\mathbf{v}$ is a sum of orthogonal projections of $\mathbf{v}$ onto the directions $\mathbf{q}_i$ (the ith projection is achieved by the special rank-one matrix $\mathbf{q}_i \mathbf{q}_i^T$ – See Lecture 6)

$\diamond$ ORTHOGONAL MATRICES

A square matrix $\mathbf{Q} \in \mathbb{R}^{m\times m}$ is *orthogonal* if

$$\mathbf{Q^T} = \mathbf{Q^{-1}}.$$

i.e.,
$$\mathbf{Q^T Q} = \mathbf{Q Q^T} = \mathbf{I}$$

$$\begin{bmatrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \\ \vdots \\ \mathbf{q}_m^T \end{bmatrix} \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_m \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

<u>NOTATION</u>

$$\mathbf{q}_i^T \mathbf{q}_j = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \qquad \begin{array}{l} \delta_{ij} \text{ is called the} \\ \textit{Kronecker delta} \end{array}$$

$\diamond$ MULTIPLICATION BY AN ORTHOGONAL MATRIX

$\mathbf{b}$: coefficients of expansion using canonical basis $\{e_1, e_2, ..., e_m\}$	$\overset{\mathbf{Q^T}}{\longrightarrow}$ $\underset{\mathbf{Q}}{\longleftarrow}$	$\mathbf{Q^T b}$: coefficients of expansion using basis of columns of $\mathbf{Q^T}$ (rows of $\mathbf{Q}$)

Note 4. *Geometric structure is preserved!*

$$(\mathbf{Qx})^T(\mathbf{Qy}) = \mathbf{x^T y} \qquad \textit{(verify!)}$$

This invariance of inner products implies angles between vectors are preserved, and so are their lengths, $\|\mathbf{Qx}\| = \|\mathbf{x}\|$.

- *Corresponds to rigid rotation of the vector space if* $\det(\mathbf{Q}) = +1$ *or reflection if* $\det(\mathbf{Q}) = -1$.