A review of CAAM 336 linear algebra

12 September 2014

Symmetric matrices show up in many numerical methods and simulation techniques engineering applications. Very often, this is due to the fact that methods and techniques are *designed* to give rise to square symmetric matrices, in order to take advantage of the wonderful mathematical properties that arise from symmetry.

We'll review here basic linear algebra concepts: vectors and matrices, and eigenvalues/eigenvectors of symmetric matrices.

Basic linear algebra

Let's define some notation:

1. A vector with N terms $\mathbf{v}$ is defined such that

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix}.$$

We'll refer to the ith term of a vector using the notation $(\mathbf{v})_i = v_i$.

2. An $N \times N$ matrix $\mathbf{A}$ is defined with entries a_{ij} such that

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \vdots & & \ddots & \\ a_{N1} & a_{N2} & \dots & a_{NN} \end{bmatrix}, \qquad \mathbf{A}_{ij} = a_{ij}.$$

We can define a vector and matrix transpose as well

1. $\mathbf{v}^T$ is defined such that

$$\mathbf{v}^T = \begin{bmatrix} v_1, & v_2, & \dots, & v_N \end{bmatrix}.$$

2. The matrix transpose $\mathbf{A}^T$ is defined entrywise

$$\mathbf{A}^T_{ij} = \mathbf{A}_{ji}$$

For example, when

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix},$$

then $\mathbf{A}^T$ is

$$\mathbf{A}^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}.$$

In other words, the entries are flipped across the diagonal.

It's helpful also to define ways to measure the magnitude of a vector; to this end, we'll define the vector norm

$$\|\mathbf{v}\| = \sqrt{\sum_{i=1}^{N} v_i}.$$

Note that, for $N = 3$, the vector norm reduces down to the usual magnitude of a vector:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}, \qquad \|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + v_3^2}.$$

We refer to a vector $\mathbf{v}$ as a *unit vector* if it has norm $\|\mathbf{v}\| = 1$.

Operations on matrices and vectors

We can define operations like multiplication involving matrices and vectors as well.

1. A dot product $\mathbf{u}^T\mathbf{v}$ between two vectors $\mathbf{u}$ and $\mathbf{v}$ is defined as

   $$\mathbf{u}^T\mathbf{v} = \sum_{i=1}^{N} u_i v_i$$

   This definition is symmetric; in other words, $\mathbf{u}^T\mathbf{v} = \mathbf{v}^T\mathbf{u}$. Notice that the vector norm $\|\mathbf{v}\|$ can be defined in terms of the dot product:

   $$\|\mathbf{v}\|^2 = \sum_{i=1}^{N} v_i^2 = \mathbf{v}^T\mathbf{v}.$$

2. Two vectors $\mathbf{u}$ and $\mathbf{v}$ are considered *orthogonal* if

   $$\mathbf{u}^T\mathbf{v} = 0.$$

3. A matrix $\mathbf{A}$ multiplied by a vector $\mathbf{v}$ gives back

   $$(\mathbf{Av})_i = \sum_{j=1}^{N} a_{ij} v_j$$

   or

   $$\mathbf{Av} = \begin{bmatrix} \sum_{j=1}^{N} a_{1j} v_j \\ \sum_{j=1}^{N} a_{2j} v_j \\ \vdots \\ \sum_{j=1}^{N} a_{Nj} v_j \end{bmatrix}.$$

   In other words, the dot product of the ith *row* of the matrix $\mathbf{A}$ and the vector $\mathbf{v}$ gives back the ith term of the matrix-vector product $\mathbf{Av}$.

For example, consider the vectors $\mathbf{u} = [1, 0]^T$ and $\mathbf{v} = [0, 1]^T$. If we plotted these vectors, they would lie on the x and y axis, respectively, and have a 90° angle between them. They are perpendicular to each other, which is analogous to two orthogonality in 2D. Checking their dot products, we can see that $\mathbf{u}^T\mathbf{v} = 0$ as well.

Linear independence

A set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is said to be *linearly dependent* if one or more vectors $\mathbf{v}_j$ can be written as a linear combination of the other vectors, or that

$$\mathbf{v}_j = \sum_{i \neq j} \alpha_i \mathbf{v}_i$$

for some scalar values α_i. If a set of vectors is not linearly dependent, we refer to it as *linearly independent*.

For example, the set of three vectors

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$$

are linearly dependent - the first two vectors are linearly independent, but the third is not, since

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

The reason this definition is useful is because a matrix $\mathbf{A}$ has an inverse $\mathbf{A}^{-1}$ such that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = I$ if and only if the columns $\mathbf{a_i}$ of the matrix

$$\mathbf{A} = \begin{bmatrix} | & | & & | \\ \mathbf{a_1} & \mathbf{a_2} & \cdots & \mathbf{a_N} \\ | & | & & | \end{bmatrix}$$

are linearly independent.

Eigenvalues and eigenvectors

Eigenvalues and eigenvectors are possibly the most important concept to permeate engineering mathematics from linear algebra. We can start with their definition:

Definition 1 *A vector* $\mathbf{v}$ *is an eigenvector with associated eigenvalue* λ *of the matrix* $\mathbf{A}$ *if*

$$\mathbf{A}\mathbf{v} = \lambda \mathbf{v}.$$

In other words, eigenvectors of $\mathbf{A}$ are vectors such that, when multiplied by $\mathbf{A}$, give back a scaling of the original vector.

Suppose we are given a matrix $\mathbf{A}$ with eigenvalues λ_i and eigenvectors $\mathbf{v}_i$. Then, if we form another matrix using $\mathbf{A}$, we can sometimes determine the eigenvalues and eigenvectors of that matrix based on the eigenvalues and eigenvectors of $\mathbf{A}$. For example, consider

$$\mathbf{B} = I + \alpha \mathbf{A}.$$

$\mathbf{B}$ has eigenvalues $1 + \alpha\lambda_i$ and eigenvectors $\mathbf{v}_i$, which we can show by noting that

$$\mathbf{B}\mathbf{v}_i = (I + \alpha\mathbf{A})\mathbf{v}_i = \mathbf{v}_i + \alpha\mathbf{A}\mathbf{v}_i = \mathbf{v}_i + \alpha\lambda_i\mathbf{v}_i = (1 + \alpha\lambda_i)\mathbf{v}_i.$$

The eigenvalue decomposition

Another fact is that all of the $N \times N$ matrices we'll deal with in this class will have N distinct eigenvalues and eigenvectors. We will index them using i, such that $\mathbf{v}_i$ and λ_i are an eigenvector/eigenvalue of $\mathbf{A}$ such that

$$\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i.$$

We can stack all the vectors $\mathbf{v}_i$ into the columns of a matrix. Multiplying this matrix by $\mathbf{A}$ is the same as multiplying each column by $\mathbf{A}$; the result is

$$\mathbf{A}\mathbf{V} = \begin{bmatrix} | & | & & | \\ \mathbf{A}\mathbf{v}_1 & \mathbf{A}\mathbf{v}_2 & \cdots & \mathbf{A}\mathbf{v}_N \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \cdots & \lambda_N\mathbf{v}_N \\ | & | & & | \end{bmatrix}.$$

Let's note that multiplication on the left by a diagonal matrix is equivalent to scaling the columns by the diagonal entries. Then, we can factor the above into

$$\begin{bmatrix} | & | & & | \\ \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \cdots & \lambda_N\mathbf{v}_N \\ | & | & & | \end{bmatrix} =$$

$$\begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_N \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \end{bmatrix},$$

or, in the more compact form,

$$\mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{\Lambda}.$$

Definition 2 *Since we've assumed all our $\mathbf{v}$'s are linearly independent from each other, $\mathbf{V}$ has linearly independent columns and is thus invertible. We can multiply both sides of the above equation by $\mathbf{V}^{-1}$ on the right to get* the eigenvalue decomposition *of $\mathbf{A}$*

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}.$$

Usefulness of the eigenvalue decomposition

The eigenvalue decomposition is particularly useful in analyzing *powers* of a matrix. Observe that

$$\mathbf{A}^2 = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}\mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1} = \mathbf{V}\mathbf{\Lambda}^2\mathbf{V}^{-1}.$$

Since $\mathbf{\Lambda}$ is diagonal,

$$\mathbf{\Lambda}^2 = \begin{bmatrix} \lambda_1^2 & & & \\ & \lambda_2^2 & & \\ & & \ddots & \\ & & & \lambda_N^2 \end{bmatrix}.$$

This generalizes to higher powers of $\mathbf{A}$ as well

$$\mathbf{A}^j = \mathbf{V\Lambda^j V^{-1}}.$$

We can also use it to characterize matrix inverses: note that

$$\mathbf{A}(\mathbf{V\Lambda^{-1}V^{-1}}) = (\mathbf{V\Lambda^{-1}V^{-1}})\mathbf{V\Lambda V^{-1}} = \mathbf{I}.$$

This implies that we can write the inverse of $\mathbf{A}$ as

$$\mathbf{A}^{-1} = \mathbf{V\Lambda^{-1}V^{-1}}.$$

The summation form of the eigenvalue decomposition

We can also express this above decomposition as a summation; define the matrix $\mathbf{W}$ to be

$$\mathbf{W} = \mathbf{V^{-1}}.$$

Let $\mathbf{w_i^T}$ be the ith row of $\mathbf{W}$; then, we can write the eigenvalue decomposition as

$$\mathbf{A} = \begin{bmatrix} | & | & & | \\ \lambda_1 \mathbf{v}_1 & \lambda_2 \mathbf{v}_2 & \cdots & \lambda_N \mathbf{v}_N \\ | & | & & | \end{bmatrix} \cdot \begin{bmatrix} \text{———} & \mathbf{w_1^T} & \text{———} \\ \text{———} & \mathbf{w_2^T} & \text{———} \\ & \cdots & \\ \text{———} & \mathbf{w_N^T} & \text{———} \end{bmatrix}$$

$$= \sum_{i=1}^{n} \lambda_i \mathbf{v}_i \mathbf{w_i^T}$$

This is called the a *dyadic* form of the eigenvalue decomposition.

Symmetric matrices

We now have enough background to present some specific facts about symmetric matrices.

Theorem 1 *All eigenvalues of a symmetric matrix are real.*

The proof can be found in most standard linear algebra textbooks, but involves some more notation than what we've introduced here. For the sake of brevity, we'll skip it.

Theorem 2 (Spectral Theorem) *Suppose* $\mathbf{A} \in \mathcal{R}^{N\times N}$ *is symmetric. Then there exist N eigenvalues* $\lambda_1, \ldots, \lambda_N$ *and corresponding unit-length eigenvectors* $\mathbf{v}_1, \ldots, \mathbf{v}_N$ *such that*

$$\mathbf{A}\mathbf{v}_j = \lambda_j \mathbf{v}_j.$$

The eigenvectors are linearly independent, and $\mathbf{v}_j^T\mathbf{v}_k = 0$ *when* $j \neq k$*, and* $\mathbf{v}_j^T\mathbf{v}_j = \|\mathbf{v}_j\|^2 = 1$.

For example, when

$$\mathbf{A} = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix},$$

we have $\lambda_1 = 4$ and $\lambda_2 = 2$, with

$$\mathbf{v}_1 = \begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}.$$

Note that these eigenvectors are unit vectors, and they are orthogonal.

Since all the eigenvectors are linearly independent, we can conclude that

$$\mathbf{V} = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_N \\ | & | & & | \end{bmatrix}$$

is invertible, and has an inverse $\mathbf{V}$. By $\mathbf{v}_j^T\mathbf{v}_k = 0$, we can also note that

$$\mathbf{V^T V} = \begin{bmatrix} \text{———} & \mathbf{v}_1^T & \text{———} \\ \text{———} & \mathbf{v}_2^T & \text{———} \\ & \cdots & \\ \text{———} & \mathbf{v}_N^T & \text{———} \end{bmatrix} \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_N \\ | & | & & | \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{v}_1^T\mathbf{v}_1 & \mathbf{v}_1^T\mathbf{v}_2 & \dots & \mathbf{v}_1^T\mathbf{v}_N \\ \mathbf{v}_2^T\mathbf{v}_1 & \mathbf{v}_2^T\mathbf{v}_2 & \cdots & \mathbf{v}_2^T\mathbf{v}_N \\ \vdots & & \ddots & \vdots \\ \mathbf{v}_N^T\mathbf{v}_1 & \mathbf{v}_N^T\mathbf{v}_2 & \cdots & \mathbf{v}_N^T\mathbf{v}_N \end{bmatrix}$$

By the fact that $\mathbf{v}_i^T\mathbf{v}_j = 0$ if $i \neq j$, most of the entries in the above matrix product are zero, and

$$\mathbf{V^T V} = \mathbf{I}$$

implying that $\mathbf{V^T} = \mathbf{V^{-1}}$.

As a consequence of the above Spectral Theorem and the dyadic form of the eigenvalue decomposition, we can write any symmetric matrix $\mathbf{A} \in \mathcal{R}^{N\times N}$ in the form

$$\mathbf{A} = \sum_{j=1}^{n} \lambda_j \mathbf{v}_j \mathbf{v}_j^T. \tag{1}$$

Writing matrices in this form allows us to express the matrix-vector product $\mathbf{Au}$ as

$$\mathbf{Au} = \sum_{j=1}^{n} \lambda_j (\mathbf{v}_j^T\mathbf{u})\mathbf{v}_j.$$

A preview of things to come

The reason we focus on the summation form of the eigenvalue decomposition above is to find parallels between matrices acting on vectors and *operators* acting on functions. A matrix can be thought of a map acting on vectors: $\mathbf{A}$ applied to a vector returns back another vector. We can generalize this to functions as well. For example, consider $u(x)$ to be a function — we can define the operator L to act on functions, such that $Lu(x)$ gives back the negative of the second derivative of $u(x)$

$$Lu(x) = -\frac{\partial^2 u(x)}{\partial x^2}.$$

In the same way that the matrix **A** has eigenvalues and eigenvectors, we will find that the operator L has eigenvalues and *eigenfunctions*. For example,

$$L \sin(k\pi x) = (k\pi)^2 \sin(k\pi x)$$

implying that for $v(x) = \sin(k\pi x)$ and $\lambda_k = (k\pi)^2$,

$$Lv_k(x) = \lambda_k v_k(x).$$

We will cover these concepts in much more detail in the next several weeks of class.