Review of basic elements from Linear Algebra

A small set of basic but important ideas from matrix algebra extend to form the basis for Math 551: dot products, orthogonality, eigenvalues and eigenvectors of square matrices, transposes (adjoint matrices). This page reviews the background you'll need.

1 The dot product

1. For real-valued vectors $\mathbf{x}, \mathbf{y}, ... \in \mathbb{R}^n$, (column vectors $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ where all entries x_j are real numbers), the **dot product** of two vectors produces a scalar value (single number). The standard definition of the dot product on $\mathbb{R}^n$ is $\mathbf{x} \cdot \mathbf{y} \equiv \sum_{j=1}^{n} x_j y_j = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$; the sum of products of corresponding entries. Example: $(1, 2, 3) \cdot (4, 5, 6) = 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 = 32$
2. The dot product $\mathbf{x} \cdot \mathbf{y}$ can also be defined in terms of matrix multiplication for the product of a row-matrix[1] times a column matrix: $\mathbf{x} \cdot \mathbf{y} \equiv \mathbf{x}^T \mathbf{y} = (x_1 \ x_2 \ \cdots \ x_n) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{j=1}^{n} x_j y_j$
3. The real dot product is commutative, $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$ for all vectors.
   One way to prove this is from the properties of transposes of products, namely $(\mathbf{ABC})^T = \mathbf{C}^T \mathbf{B}^T \mathbf{A}^T$ for all possible matrices. Then to see that $\mathbf{y} \cdot \mathbf{x} = \mathbf{y}^T \mathbf{x}$ equals $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y}$ use the fact that scalars are unchanged by the transpose, $c^T = c$, so $\mathbf{x}^T \mathbf{y} = (\mathbf{x}^T \mathbf{y})^T = \mathbf{y}^T \mathbf{x}$
4. The magnitude (length, or "norm") of a vector is defined by the dot product of the vector with itself, $|\mathbf{x}|^2 = \mathbf{x} \cdot \mathbf{x} \geq 0$. The norm is zero if and only if all entries of the vector are zero, $|\mathbf{x}| = 0 \iff \mathbf{x} = \mathbf{0}$.
5. The dot product is used to define when two vectors are **orthogonal** (perpendicular) to each other: $\boxed{\mathbf{x} \perp \mathbf{y} \text{ if } \mathbf{x} \cdot \mathbf{y} = 0}$. (This is a special case from $\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}||\mathbf{y}| \cos \theta$.)

2 Eigenvalues and eigenvectors of matrices

1. If $\mathbf{A}$ is a square $n \times n$ real matrix, then the vector $\mathbf{y}$ resulting from the matrix-vector product $\mathbf{y} = \mathbf{Ax}$ can be geometrically described in general: $\mathbf{y}$ can have a different direction (check $\mathbf{x} \cdot \mathbf{y}$) and different length (check $|\mathbf{y}|$) than $\mathbf{x}$ does.
   Namely, for general vectors $\mathbf{x}$, the product $\mathbf{Ax}$ is a stretched and rotated vector.
2. For each matrix $\mathbf{A}$ there will be special choices for $\mathbf{x}$ called **eigenvectors** whose resulting $\mathbf{y}$ are **parallel** to $\mathbf{x}$. Two vectors are parallel to each other if one is a scalar multiple of the other, $\mathbf{y} = c\mathbf{x}$. Different values of the scaling constant have geometric interpretations: $c > 1$: stretched length, $0 < c < 1$: reduced length, and $c < 0$: parallel but in the opposite direction ("anti-parallel").
   When the product $\mathbf{Ax}$ is scaled but *not rotated* relative to the original $\mathbf{x}$, then that direction is an eigenvector (notation: $\boldsymbol{\phi}$ vectors). This geometric property has very important consequences for solving many problems.
3. For eigenvectors $\boldsymbol{\phi}$, since the length of $\mathbf{A}\boldsymbol{\phi}$ may be different than $\boldsymbol{\phi}$, the **eigenvalue** λ gives the scaling constant (the "c" from above). Each eigenvalue λ and its associated eigenvector $\boldsymbol{\phi}$ satisfy the equation
$$\boxed{\mathbf{A}\boldsymbol{\phi} = \lambda\boldsymbol{\phi}} \tag{1}$$
   where $\boldsymbol{\phi}$ must be a nonzero vector ($|\boldsymbol{\phi}| \neq 0$). Using the identity matrix $\mathbf{I}$, this equation can be re-written as
$$\mathbf{A}\boldsymbol{\phi} = (\lambda\mathbf{I})\boldsymbol{\phi} \qquad \to \qquad (\mathbf{A} - \lambda\mathbf{I})\boldsymbol{\phi} = \mathbf{0} \tag{2}$$

[1] a row-vector (matrix) is the transpose of a column-vector (matrix)

Recall from linear algebra that if $\mathbf{Mz} = \mathbf{0}$ with $\mathbf{z} \neq \mathbf{0}$ then matrix $\mathbf{M}$ must be singular and its determinant must be zero; this leads to the determinant equation for the eigenvalues of $\mathbf{A}$:

$$\boxed{\det(\mathbf{A} - \lambda\mathbf{I}) = 0} \tag{3}$$

Expanding out this determinant yields the characteristic polynomial $p(\lambda)$; the zeros of the characteristic polynomial, $p(\lambda) = 0$, are the eigenvalues, $\lambda_k, k = 1, 2, \cdots, n$.

4. For matrices, you start by calculating the eigenvalues from the characteristic polynomial, then you can determine the eigenvector for each eigenvalue, one at a time. If λ_k is one of the eigenvalues, then write the matrix $\mathbf{M}_k = \mathbf{A} - \lambda_k\mathbf{I}$, and do Gaussian elimination to row echelon form to determine the eigenvector $\boldsymbol{\phi}_k$ of $\mathbf{M}_k\boldsymbol{\phi}_k = \mathbf{0}$. Any multiple of an eigenvector is still the same eigenvector – length doesnt matter, only direction; you can rescale to get simpler convenient values for the entries.

5. The usual case for a $n \times n$ matrix is that you'll be able to find n different eigenvectors. This is definitely true when all of the eigenvalues are distinct. This is called the "non-defective" case, and it means the matrix has a full set of n eigenvectors (called a **complete** set) that can be used to solve every problem for the matrix. This is the case we will focus on.

2.1 Example calculation

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & -4 \\ 1 & -4 & 1 \\ -4 & 0 & 2 \end{pmatrix}$$

First, the calculate the eigenvalues: determinant eqn $\to$ characteristic polynomial $\to$ find the roots

$$\begin{aligned} \det(\mathbf{A} - \lambda\mathbf{I}) &= \begin{vmatrix} 2-\lambda & 0 & -4 \\ 1 & -4-\lambda & 1 \\ -4 & 0 & 2-\lambda \end{vmatrix} \\ &= \lambda^3 - 28\lambda - 48 \\ &= (\lambda+4)(\lambda-6)(\lambda+2) = 0 \end{aligned}$$

Then, for each λ_k, calculate the eigenvector: find a $\mathbf{x}$ to give $(\mathbf{A} - \lambda_k\mathbf{I})\mathbf{x} = \mathbf{0}$

1. For eigenvalue $\lambda_1 = -4$:

$$\mathbf{M}_1 = \mathbf{A} - \lambda_1\mathbf{I} = \begin{pmatrix} 6 & 0 & -4 \\ 1 & 0 & 1 \\ -4 & 0 & 6 \end{pmatrix}$$

Let $\mathbf{x} = (x_1, x_2, x_3)^T$. The equations for $\mathbf{M}_1\mathbf{x} = \mathbf{0}$ are

$$6x_1 - 4x_3 = 0 \qquad x_1 + x_3 = 0 \qquad -4x_1 + 6x_3 = 0$$

Doing the algebra shows that $x_1 = 0, x_3 = 0$ but x_2 can be anything (can pick $x_2 = 1$). So the scaled eigenvector is

$$\boldsymbol{\phi}_1 = (0, 1, 0)^T \qquad \mathbf{A}\boldsymbol{\phi}_1 = -4\boldsymbol{\phi}_1$$

2. For eigenvalue $\lambda_2 = 6$:

$$\mathbf{M}_2 = \mathbf{A} - \lambda_2\mathbf{I} = \begin{pmatrix} -4 & 0 & -4 \\ 1 & -10 & 1 \\ -4 & 0 & -4 \end{pmatrix}$$

Equations for $\mathbf{M}_2\mathbf{x} = \mathbf{0}$:

$$-4x_1 - 4x_3 = 0 \qquad x_1 - 10x_2 + x_3 = 0 \qquad -4x_1 - 4x_3 = 0$$

Need $x_1 = -x_3$ and $x_2 = 0$ but x_3 can be anything (can pick $x_3 = 1$). So the scaled eigenvector is

$$\boldsymbol{\phi}_2 = (1, 0. -1)^T \qquad \mathbf{A}\boldsymbol{\phi}_2 = 6\boldsymbol{\phi}_2$$

3. For eigenvalue $\lambda_3 = -2$:

$$\mathbf{M}_3 = \mathbf{A} - \lambda_3\mathbf{I} = \begin{pmatrix} 4 & 0 & -4 \\ 1 & -2 & 1 \\ -4 & 0 & 4 \end{pmatrix}$$

Equations for $\mathbf{M}_3\mathbf{x} = \mathbf{0}$:

$$4x_1 - 4x_3 = 0 \qquad x_1 - 2x_2 + x_3 = 0 \qquad -4x_1 + 4x_3 = 0$$

Need $x_1 = x_3$ and $x_2 = x_3$ but x_3 can be anything (can pick $x_3 = 1$). So the scaled eigenvector is

$$\boldsymbol{\phi}_3 = (1, 1, 1)^T \qquad \mathbf{A}\boldsymbol{\phi}_3 = -2\boldsymbol{\phi}_3$$

This matrix is not symmetric ($\mathbf{A}^T \neq \mathbf{A}$) so it will have different adjoint eigenvectors, $\{\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \boldsymbol{\psi}_3\}$ than you can calculate similarly using $\mathbf{M}_k = \mathbf{A}^T - \lambda_k\mathbf{I}$. These will be needed for calculating the c_k's to solve $\mathbf{Au} = \mathbf{b}$ as $\mathbf{u} = c_1\boldsymbol{\phi}_1 + c_2\boldsymbol{\phi}_2 + c_3\boldsymbol{\phi}_3$.

For more information, see any basic textbook on linear algebra, for example, *Introduction to Linear Algebra* or *Linear algebra and its applications* by Gilbert Strang.