LINEAR ALGEBRA. LECTURE 10

Eigenvalues and eigenvectors

The subject of eigenvalues and eigenvectors will take up most of the rest of the course. We will again be working with square matrices. Eigenvalues are special numbers associated with a matrix and eigenvectors are special vectors.

Eigenvectors and eigenvalues

A matrix A *acts on* vectors $\mathbf{x}$ like a function does, with input $\mathbf{x}$ and output $A\mathbf{x}$. *Eigenvectors* are vectors for which $A\mathbf{x}$ is parallel to $\mathbf{x}$. In other words:

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

In this equation, $\mathbf{x}$ is an eigenvector of A and λ is an *eigenvalue* of A.

Eigenvalue 0

If the eigenvalue λ equals 0 then $A\mathbf{x} = 0\mathbf{x} = \mathbf{0}$. Vectors with eigenvalue 0 make up the nullspace of A; if A is singular, then $\lambda = 0$ is an eigenvalue of A.

Examples

Suppose P is the matrix of a projection onto a plane. For any $\mathbf{x}$ in the plane $P\mathbf{x} = \mathbf{x}$, so $\mathbf{x}$ is an eigenvector with eigenvalue 1. A vector $\mathbf{x}$ perpendicular to the plane has $P\mathbf{x} = \mathbf{0}$, so this is an eigenvector with eigenvalue $\lambda = 0$. The eigenvectors of P span the whole space (but this is not true for every matrix).

The matrix $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ has an eigenvector $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ with eigenvalue 1 and another eigenvector $\mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ with eigenvalue -1. These eigenvectors span the space. They are perpendicular because $B = B^T$ (as we will prove).

$\det(A - \lambda I) = 0$

An n by n matrix will have n eigenvalues, and their sum will be the sum of the diagonal entries of the matrix: $a_{11} + a_{22} + \cdots + a_{nn}$. This sum is the *trace* of the matrix. For a two by two matrix, if we know one eigenvalue we can use this fact to find the second.

Can we solve $A\mathbf{x} = \lambda\mathbf{x}$ for the eigenvalues and eigenvectors of A? Both λ and $\mathbf{x}$ are unknown; we need to be clever to solve this problem:

$$\begin{aligned} A\mathbf{x} &= \lambda\mathbf{x} \\ (A - \lambda I)\mathbf{x} &= \mathbf{0} \end{aligned}$$

In order for λ to be an eigenvector, $A - \lambda I$ must be singular. In other words, $\det(A - \lambda I) = 0$. We can solve this *characteristic equation* for λ to get n solutions.

If we're lucky, the solutions are distinct. If not, we have one or more *repeated eigenvalues*.

Once we've found an eigenvalue λ, we can use elimination to find the nullspace of $A - \lambda I$. The vectors in that nullspace are eigenvectors of A with eigenvalue λ.

Calculating eigenvalues and eigenvectors

Let $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. Then:

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{vmatrix} \\ &= (3-\lambda)^2 - 1 \\ &= \lambda^2 - 6\lambda + 8. \end{aligned}$$

Note that the coefficient 6 is the trace (sum of diagonal entries) and 8 is the determinant of A. In general, the eigenvalues of a two by two matrix are the solutions to:

$$\lambda^2 - \text{trace}(A) \cdot \lambda + \det A = 0.$$

Just as the trace is the sum of the eigenvalues of a matrix, the product of the eigenvalues of any matrix equals its determinant.

For $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$, the eigenvalues are $\lambda_1 = 4$ and $\lambda_2 = 2$. We find the eigenvector $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ for $\lambda_1 = 4$ in the nullspace of $A - \lambda_1 I = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$.

x_2 will be in the nullspace of $A - 2I = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. The nullspace is an entire line; x_2 could be any vector on that line. A natural choice is $x_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Note that these eigenvectors are the same as those of $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Adding $3I$ to the matrix $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ added 3 to each of its eigenvalues and did not change its eigenvectors, because $A\mathbf{x} = (B + 3I)\mathbf{x} = \lambda\mathbf{x} + 3\mathbf{x} = (\lambda + 3)\mathbf{x}$.

A caution

Similarly, if $A\mathbf{x} = \lambda\mathbf{x}$ and $B\mathbf{x} = \alpha\mathbf{x}$, $(A + B)\mathbf{x} = (\lambda + \alpha)\mathbf{x}$. It would be nice if the eigenvalues of a matrix sum were always the sums of the eigenvalues, but this is only true if A and B have the same eigenvectors. The eigenvalues of the product AB aren't usually equal to the products $\lambda(A)\lambda(b)$, either.

Complex eigenvalues

The matrix $Q = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ rotates every vector in the plane by 90°. It has trace $0 = \lambda_1 + \lambda_2$ and determinant $1 = \lambda_1 \cdot \lambda_2$. Its only real eigenvector is the zero vector; any other vector's direction changes when it is multiplied by Q. How will this affect our eigenvalue calculation?

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} \\ &= \lambda^2 + 1. \end{aligned}$$

$\det(A - \lambda I) = 0$ has solutions $\lambda_1 = i$ and $\lambda_2 = -i$. If a matrix has a complex eigenvalue $a + bi$ then the *complex conjugate* $a - bi$ is also an eigenvalue of that matrix.

Symmetric matrices have real eigenvalues. For *antisymmetric* matrices like Q, for which $A^T = -A$, all eigenvalues are imaginary ($\lambda = bi$).

Triangular matrices and repeated eigenvalues

For triangular matrices such as $A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$, the eigenvalues are exactly the entries on the diagonal. In this case, the eigenvalues are 3 and 3:

$$\begin{aligned} \det(A - \lambda \det I) &= \begin{vmatrix} 3-\lambda & 1 \\ 0 & 3-\lambda \end{vmatrix} \\ &= (3-\lambda)(3-\lambda) \quad \Big(= (a_{11} - \lambda)(a_{22} - \lambda) \Big) \\ &= 0, \end{aligned}$$

so $\lambda_1 = 3$ and $\lambda_2 = 3$. To find the eigenvectors, solve:

$$(A - \lambda I)\mathbf{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{x} = \mathbf{0}$$

to get $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. There is no independent eigenvector $\mathbf{x}_2$.

Diagonalization and powers of A

We know how to find eigenvalues and eigenvectors. In this lecture we learn to *diagonalize* any matrix that has n independent eigenvectors and see how diagonalization simplifies calculations. The lecture concludes by using eigenvalues and eigenvectors to solve *difference equations*.

Diagonalizing a matrix $S^{-1}AS = \Lambda$

If A has n linearly independent eigenvectors, we can put those vectors in the columns of a (square, invertible) matrix S. Then

$$\begin{aligned} AS &= A \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1\mathbf{x}_1 & \lambda_2\mathbf{x}_2 & \cdots & \lambda_n\mathbf{x}_n \end{bmatrix} \\ &= S \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} = S\Lambda. \end{aligned}$$

Note that Λ is a diagonal matrix whose non-zero entries are the eigenvalues of A. Because the columns of S are independent, S^{-1} exists and we can multiply both sides of $AS = S\Lambda$ by S^{-1}:

$$S^{-1}AS = \Lambda.$$

Equivalently, $A = S\Lambda S^{-1}$.

Powers of A

What are the eigenvalues and eigenvectors of A^2?

$$\begin{aligned} \text{If } A\mathbf{x} &= \lambda\mathbf{x}, \\ \text{then } A^2\mathbf{x} &= \lambda A\mathbf{x} = \lambda^2\mathbf{x}. \end{aligned}$$

The eigenvalues of A^2 are the squares of the eigenvalues of A. The eigenvectors of A^2 are the same as the eigenvectors of A. If we write $A = S\Lambda S^{-1}$ then:

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

Similarly, $A^k = S\Lambda^k S^{-1}$ tells us that raising the eigenvalues of A to the kth power gives us the eigenvalues of A^k, and that the eigenvectors of A^k are the same as those of A.

Theorem: If A has n independent eigenvectors with eigenvalues λ_i, then $A^k \to 0$ as $k \to \infty$ if and only if all $|\lambda_i| < 1$.

A is guaranteed to have n independent eigenvectors (and be *diagonalizable*) if all its eigenvalues are different. Most matrices do have distinct eigenvalues.

Repeated eigenvalues

If A has repeated eigenvalues, it may or may not have n independent eigenvectors. For example, the eigenvalues of the identity matrix are all 1, but that matrix still has n independent eigenvectors.

If A is the triangular matrix $\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$ its eigenvalues are 2 and 2. Its eigenvectors are in the nullspace of $A - \lambda I = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ which is spanned by $\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. This particular A does not have two independent eigenvectors.

Difference equations $\mathbf{u}_{k+1} = A\mathbf{u}_k$

Start with a given vector $\mathbf{u}_0$. We can create a sequence of vectors in which each new vector is A times the previous vector: $\mathbf{u}_{k+1} = A\mathbf{u}_k$. $\mathbf{u}_{k+1} = A\mathbf{u}_k$ is a *first order difference equation*, and $\mathbf{u}_k = A^k\mathbf{u}_0$ is a solution to this system.

We get a more satisfying solution if we write $\mathbf{u}_0$ as a combination of eigenvectors of A:

$$\mathbf{u}_0 = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n = S\mathbf{c}.$$

Then:

$$A\mathbf{u}_0 = c_1\lambda_1\mathbf{x}_1 + c_2\lambda_2\mathbf{x}_2 + \cdots + c_n\lambda_n\mathbf{x}_n$$

and:

$$\mathbf{u}_k = A^k\mathbf{u}_0 = c_1\lambda_1^k\mathbf{x}_1 + c_2\lambda_2^k\mathbf{x}_2 + \cdots + c_n\lambda_n^k\mathbf{x}_n = S\Lambda^k\mathbf{c}.$$

Simple iteration method

New solution methods are needed when a problem $Ax = b$ is too large and expensive for ordinary elimination. We are thinking of *sparse matrices A*, so that multiplications Ax are relatively cheap. If A has at most p nonzeros in every row, then *Ax needs at most pn multiplications*.

We are turning from elimination to look at iterative methods. The goal of numerical linear algebra is clear: Find a fast stable algorithm that uses the special properties of the matrix. We meet matrices that are symmetric or triangular or orthogonal or tridiagonal. Those are at the core of matrix computations. The algorithm doesn't need details of the entries (which come from the specific application). By concentrating on the matrix structure, numerical linear algebra offers major help.

Overall, elimination with good numbering is the first choice! But storage and CPU time can become excessive, especially in three dimensions. At that point we turn from elimination to iterative methods. Get a new form of $Ax = b$:

$$Ax = b \iff x + \tau Ax = \tau b + x \iff$$
$$x = (I - \tau A)x + \tau b = Bx + \tau b$$

were $B = (I - \tau A)$ and τ is an iteration parameter. This form suggests an iteration, in which every vector x_k leads to the next x_{k+1}:

$$x_{k+1} = (I - \tau A)x_k + \tau b = Bx_k + \tau b$$

Iteration starting from any x_0, the first step finds x_1 from $x_1 = Bx_0 + \tau b$ and so on. Two conditions on B make the iteration successful:

1) the new x_{k+1} must be quickly computable. Equation $x_{k+1} = Bx_k + \tau b$ must be fast to solve;

2) the error $\varepsilon_{k+1} = x_{k+1} - x$ should approach zero as rapidly as possible.

Subtract first equation from second to find the error equation. It connects ε_k to ε_{k+1}: Error $\varepsilon_{k+1} = B\varepsilon_k = (I - \tau A)\varepsilon_k$. The right side b disappears in this error equation. Each step multiplies the error vector ε_k by B. The speed of convergence of x_k to x (and ε_k to zero) depends entirely on B. *The test for convergence is given by the eigenvalues of B*:

Convergence test	Every eigenvalue of $B = I - \tau A$ must have $\lvert\lambda\rvert < 1$.

The largest eigenvalue (in absolute value) is the spectral radius $\rho(B) = \max\lvert\lambda(B)\rvert$. Convergence requires $\rho(B) < 1$. The convergence rate is set by the largest eigenvalue. To get this condition we can use our iteration parameter τ .