MA-106 Linear Algebra

M.K. Keshari

Department of Mathematics
Indian Institute of Technology Bombay
Powai, Mumbai - 76

17th February, 2015
D1 - Lecture 19

Recall:

- A square matrix A is diagonalizable if and only if A has a basis consisting of eigenvectors.
- For an invertible matrix S, $S^{-1}AS$ is a diagonal matrix if and only columns of S are eigenvectors of A.
- If $x_1, \ldots, x_r$ are eigenvectors of A associated to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$, then $x_1, \ldots, x_r$ are linearly independent.
- If A has n distinct eigenvalues, then A is diagonalizable.
- A triangular matrix need not be diagonalizable.
- Assume A and B are diagonalizable . Then A and B have same set of eigenvectors if and only if $AB = BA$.

Eigenvalues of A^k

- If $Ax = \lambda x$, then $A^2x = A(Ax) = A(\lambda x) = \lambda(Ax) = \lambda^2 x$.

  Similarly $A^k x = \lambda^k x$ for any $k \geq 0$.

  Thus if λ is an eigenvalue of A with associated eigenvector x, then $\boxed{\lambda^k \text{ is an eigenvalue for } A^k\text{, with the same associated eigenvector } x}$ for $k \geq 0$. If A is invertible, then $\lambda \neq 0$. Hence, the same also holds for $k < 0$ since $A^{-1}x = \lambda^{-1}x$.

- Assume $\boxed{A \text{ is diagonalizable}}$. Then $S^{-1}AS = \Lambda$ is diagonal where columns of S are eigenvectors of A.

  Now $(S^{-1}AS)(S^{-1}AS) = S^{-1}A^2S = \Lambda^2$, which is diagonal. Hence eigenvectors of A^2 are same as eigenvectors of A. Similarly $(S^{-1}AS)^k = S^{-1}A^kS = \Lambda^k$. $\boxed{\text{For } k \geq 0,\ A^k \text{ is diagonalizable}}$ and the same also holds for $k < 0$ if A is invertible.

Application: Fibonacci Numbers

Let $F_0 = 0$, $F_1 = 1$ and $F_k = F_{k-1} + F_{k-2}$ for $k \geq 2$ define the Fibonacci sequence. What is the kth term?

If $u_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}$, then $\begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_k \\ F_{k-1} \end{bmatrix}$, i.e.,

$u_k = Au_{k-1}$ for $k \geq 1$, where $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \Rightarrow u_k = A^k u_0$ for $k \geq 1$.

Characteristic polynomial of A: $(1-\lambda)(-\lambda) - 1 = \lambda^2 - \lambda - 1$.

Eigenvalues: $\lambda_1 = \dfrac{1+\sqrt{5}}{2}$, $\lambda_2 = \dfrac{1-\sqrt{5}}{2}$.

The eigenvalues are distinct $\Rightarrow$ the associated eigenvectors x_1 and x_2 are linearly independent $\Rightarrow \{x_1, x_2\}$ is a basis for $\mathbb{R}^2$.

Write $u_0 = c_1x_1 + c_2x_2$. Then $u_k = A^k u_0 = A^k(c_1x_1 + c_2x_2)$

$$= c_1A^kx_1 + c_2A^kx_2 = c_1\left(\frac{1+\sqrt{5}}{2}\right)^k x_1 + c_2\left(\frac{1-\sqrt{5}}{2}\right)^k x_2.$$

Q: Find x_1, x_2, c_1 and c_2 and get the exact formula for F_k.

An Application: Steady State

Suppose we have a system where the current state depends linearly on the previous one, i.e., $u_k = Au_{k-1}$. Then observe that $u_k = A^k u_0$.
The **steady state** of the system is $u_\infty = \lim_{k\to\infty}(u_k)$.
Q. How do we find the steady state?

- If u_0 is an eigenvector of A associated to λ, then $u_k = \lambda^k u_0$.
- If $x_1, \ldots, x_r$ are eigenvectors of A associated to $\lambda_1, \ldots, \lambda_r$ resp. and $u_0 \in \operatorname{Span}\{x_1, \ldots, x_r\}$, i.e., $u_0 = c_1x_1 + \cdots + c_rx_r$ for scalars $c_1, \ldots, c_r$, then
  $u_k = A^k u_0 = c_1A^kx_1 + \cdots + c_rA^kx_r = c_1\lambda_1^kx_1 + \cdots + c_r\lambda_r^kx_r$.
- If A is diagonalizable, then there is a basis of $\mathbb{R}^n$ consisting of eigenvectors of A. Hence, the previous remark holds for all $u_0 \in \mathbb{R}^n$.
- Let A be diagonalizable and u_k represent population of some species.
  Under what conditions will there be a population explosion?
  What conditions will force the population to become extinct?
  When does the population stabilise to a non-zero value?
  Hint: Answer depends on eigenvalues λ_i of A.

An Application: Predator - Prey Model

Let us study the dynamics of spotted owl and wood rat polpulation. Assume owls diet are mostly rats and likewise rats are mostly predated by owls. Let O_k and R_k denote population of owls and rats in month k. Let us model the population dynamics as follows

$$\begin{aligned} O_{k+1} &= 0.5O_k + 0.4R_k \\ R_{k+1} &= -pO_k + 1.1R_k \end{aligned}$$

Here pO_k is the no. of rats consumed per month per owl.
1.1 is growth rate of rats in the absence of owls, (10% increase monthly)
$0.5O_k$ means 50% decrease in owls population if there are no rats.
$0.4R_k$ means growth rate of owl population if rats are present.
Write the system as $u_{k+1} = Au_k$, where $u_k = \begin{pmatrix} O_k \\ R_k \end{pmatrix}$, $A = \begin{pmatrix} 0.5 & 0.4 \\ -p & 1.1 \end{pmatrix}$.
Assume $u_0 = (100 \quad 100)^T$ and $p = 0.2$. Then $u_k = A^k u_0$.
Find whether the populations grow, decline or approach a steady state.

Complex Eigenvalues

Ex: Rotation by 90^o, $K = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, $|K - \lambda I| = \lambda^2 + 1$.

It has no eigenvector, since rotation by 90^o changes the direction.

K has eigenvalues, but they are not real. They are imaginary numbers i and $-i$, where $i^2 = -1$.
The associated eigenvectors are also not real. Let us compute them.

$$(K - iI)x_1 = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} x_1 = 0 \Rightarrow x_1 = \begin{bmatrix} 1 \\ -i \end{bmatrix},$$

$$(K + iI)x_2 = \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} x_2 = 0 \Rightarrow x_2 = \begin{bmatrix} 1 \\ i \end{bmatrix}.$$

The eigenvalues are distinct (though imaginary), hence eigenvectors are linearly independent.

If $S = \begin{bmatrix} x_1 & x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}$, then $S^{-1}KS = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$.

Complex Numbers

Conclusion: We need complex numbers $\mathbb{C}$ even if we are working with real matrices. Over $\mathbb{C}$, an $n \times n$ matrix A always has n eigenvalues.

Reason: $\boxed{\text{Fundamental theorem of Algebra}}$

Every polynomial over $\mathbb{C}$ of degree n has n roots in $\mathbb{C}$.

For a complex number $x = a + ib$, its conjugate is $\bar{x} = a - ib$.
Note that $\bar{x} \cdot x = a^2 + b^2$.
If $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbb{C}^n$ are complex vectors,
define $\boxed{x \cdot y = \bar{x_1} y_1 + \cdots + \bar{x_n} y_n} \in \mathbb{C}$. Then

- $x \cdot x = |x_1|^2 + \cdots + |x_n|^2 \in \mathbb{R}$ and is ≥ 0.
- $x \cdot x = 0 \Leftrightarrow x = 0$.
- For $c \in \mathbb{C}$, $x \cdot c\, x = \bar{x_1} c\, x_1 + \cdots + \bar{x_n} c\, x_n = (x \cdot x)c$.
- Observe $c\, x \cdot x = \bar{c}(x \cdot x)$

(In fact, this defines a complex inner product on $\mathbb{C}^n$).

Real Symmetric $n \times n$ Matrix A:

- For every $x \in \mathbb{C}^n$, $\boxed{x \cdot Ax \in \mathbb{R}}$. Let us do it for $n = 2$.

  If $x = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{C}^2$, $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$, then $Ax = \begin{pmatrix} az_1 + bz_2 \\ bz_1 + dz_2 \end{pmatrix}$.

  Now $x \cdot Ax = \bar{z}_1(az_1 + bz_2) + \bar{z}_2(bz_1 + dz_2)$
  $$= (a|z_1|^2 + d|z_2|^2) + b(\bar{z}_1 z_2 + \bar{z}_2 z_1) \in \mathbb{R}.$$

- $\boxed{\text{Every eigenvalue of } A \text{ is real}}$.

  Let $\lambda \in \mathbb{C}$ be an eigenvalue of A, with associated eigenvector $x \in \mathbb{C}^n$.
  Then $Ax = \lambda x \Rightarrow x \cdot Ax = x \cdot \lambda x = (x \cdot x)\lambda$.

  Since $x \neq 0 \Rightarrow x \cdot x$ is a positive real number $\Rightarrow \lambda = \dfrac{x \cdot Ax}{x \cdot x} \in \mathbb{R}$.

- A has real eigenvector associated to (real) eigenvalue λ.

  This follows from the fact that dim $N(A - \lambda I) \geq 1$.

  We can also find it from complex eigenvector $x \in \mathbb{C}^n$.
  Write $x = u + iv$, where u, $v \in \mathbb{R}^n$. Then $Ax = \lambda x \Rightarrow$
  $Au + iAv = \lambda(u + iv) \Rightarrow Au = \lambda u$ and $Av = \lambda v$.
  Therefore u and v are eigenvector of A associated to λ.

Real Symmetric $n \times n$ Matrix A:

- If $\lambda_1 \neq \lambda_2$ are distinct eigenvalues of A with eigenvectors $x_1, x_2 \in \mathbb{R}^n$, then x_1 and x_2 are orthogonal, i.e., $x_1^T x_2 = 0$.

  Now $\lambda_1(x_1^T x_2) = (\lambda_1 x_1)^T x_2 = (Ax_1)^T x_2 = (x_1^T A^T)x_2$
  $$= x_1^T(Ax_2) = x_1^T(\lambda_2 x_2) = \lambda_2(x_1^T x_2).$$
  Since $\lambda_1 \neq \lambda_2$, we get $x_1^T x_2 = 0$.

- If A has n distinct eigenvalues, then there is an orthogonal matrix Q and a diagonal matrix Λ such that $A = Q\Lambda Q^T$.

  A has n distinct eigenvalues $\Rightarrow$ A is diagonalizable $\Rightarrow S^{-1}AS = \Lambda$, where S is (real) invertible and Λ is diagonal.

  If $S = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}$ with $x_i \in \mathbb{R}^n$, then x_i's are eigenvectors of A.

  Since $x_i^T x_j = 0$ for $i \neq j$, $Q = \begin{bmatrix} \frac{x_1}{\|x_1\|} & \cdots & \frac{x_n}{\|x_n\|} \end{bmatrix}$
  is an orthogonal matrix, i.e., $Q^{-1} = Q^T$.

  Note $Q = SD$, where D is a diagonal matrix with diagonal entries $1/\|x_1\|, \dots, 1/\|x_n\|$.

  Therefore $Q^{-1}AQ = D^{-1}S^{-1}ASD = D^{-1}\Lambda D = \Lambda \Rightarrow \boxed{A = Q\Lambda Q^T}$.

Spectral Theorem for a Real Symmetric Matrix

Theorem $\boxed{\text{Every real symmetric matrix } A \text{ can be diagonalised}}$.
Further there is an orthogonal matrix Q and a diagonal matrix Λ such that $\boxed{A = Q\Lambda Q^T}$.

Once you know that A can be diagonalised, $A = Q\Lambda Q^T$ follows from the previous slide + Gram-Schmidt process applied to each eigenspace of A. (Without Proof)

Ex: Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$. Then $A - \lambda I = \begin{bmatrix} 1-\lambda & 0 & 0 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{bmatrix}$ and

$\det(A - \lambda I) = (1-\lambda)(\lambda^2 - 1)$. Eigenvalues: $\lambda_1 = 1, \lambda_2 = 1, \lambda_3 = -1$.
Eigenvectors: $x_1 = (1,0,0)^T$, $x_2 = (0,1,1)^T$ and $x_3 = (0,1,-1)$. Observe $x_1^T x_3 = 0 = x_2^T x_3$.

Gram-Schimdt gives an orthonormal basis $\{u_1, u_2\}$ for $N(A - I)$, and let $u_3 = x_3/\|x_3\|$. Then $Q = [u_1 \;\; u_2 \;\; u_3]$ is an orthogonal matrix.

Further, $A = Q\Lambda Q^T$, where Λ is a diagonal matrix with entries 1, 1, -1 on the diagonal.

For extra reading: Sketch of proof of Spectral Theorem

Since A is real symmetric, it has a real eigenvalue, say λ_1. Let x_1 be a real eigenvector of A associated to λ_1.

Let W_1 be the orthogonal complement of Span(x_1). Then $A : \mathbb{R}^n \to \mathbb{R}^n$ takes W_1 to W_1, i.e. $A|_{W_1} : W_1 \to W_1$. (Needs checking)

Since W_1 is $n-1$-dimensional, the matrix associated to $A|_{W_1}$ with respect to some basis B of W_1, i.e. $[A|_{W_1}]_B^B$, is square of size $n-1$ and symmetric.

Using induction on n, we get a basis $x_2, \ldots, x_n$ of W_1 consisting of eigenvectors of $A|_{W_1}$. Then $x_1, \ldots, x_n$ is a basis of $\mathbb{R}^n$ consisting of eigenvectors of A. This proves that A is diagonalizable.

Let $\lambda_1, \ldots, \lambda_r$ be distinct eigenvalues of A with $U_1, \ldots, U_r$ as their eigenspaces. Then U_i is orthogonal to U_j for $i \neq j$, since their elements are eigenvectors corresponding to different eigenvalues.
$\dim(U_i) =$ the number of times λ_i is repeating as eigenvalue of A.

Gram-Schmidt gives an orthonormal basis B_i of U_i.

For extra reading

Let $B = \cup_{i=1}^r B_i = \{y_1, \ldots, y_n\}$ be an orthonormal basis of $\mathbb{R}^n$. Each y_i is an eigenvector of A.

Let $Q = [y_1 \;\; y_2 \;\; \ldots \;\; Y_n]$ be an orthogonal matrix. Then $Q^{-1}AQ = \Lambda$ will be diagonal matrix. Hence $A = Q\Lambda Q^{-1}$, and $Q^{-1} = Q^T$. This completes the proof. $\square$

• Let A be symmetric, and $A = Q\Lambda Q^T$.

If $Q = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix}$ then

$$A = Q\Lambda Q^T = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix}^T$$

$$= \begin{bmatrix} \lambda_1 x_1 & \ldots & \lambda_n x_n \end{bmatrix} \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix}^T = \lambda_1 x_1 x_1^T + \ldots + \lambda_n x_n x_n^T.$$

Each matrix $\lambda_i x_i x_i^T$ is a projection matrix (check!) projecting a vector on the line spanned by x_i. Hence every symmetric matrix $\boxed{A \text{ can be decomposed as sum of } n \text{ projection matrices}}$.