Mathematics, Pusan National University

LINEAR ALGEBRA AND LEARNING FROM DATA

I.6 Eigenvalues and Eigenvectors

Taehyeong Kim
th_kim@pusan.ac.kr

July 29, 2020

Content

The eigenvectors of A don't change direction when you multiply them by A. The output $A\mathbf{x}$ is on the same line as the input vector $\mathbf{x}$. The eigenvector $\mathbf{x}$ is just multiplied by its eigenvalue λ.
If eignevectors of A $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are linearly independent, every $v \in \mathbb{R}^n$ can be expressed as

$$\begin{aligned} \mathbf{v} &= c_1\mathbf{x}_1 + \cdots + c_n\mathbf{x}_n \\ A\mathbf{v} &= c_1\lambda_1\mathbf{x}_1 + \cdots + c_n\lambda_n\mathbf{x}_n \\ A^k\mathbf{v} &= c_1\lambda_1^k\mathbf{x}_1 + \cdots + c_n\lambda_n^k\mathbf{x}_n \end{aligned}$$

Example

The rotation $Q = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ has imaginary eigenvalues $i, -i$.

$Q\begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ -i \end{bmatrix} = (i)\begin{bmatrix} 1 \\ -i \end{bmatrix}$ and $Q\begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ i \end{bmatrix} = (-i)\begin{bmatrix} 1 \\ i \end{bmatrix}$

Example

The rotation $Q = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ has imaginary eigenvalues $i, -i$.

$$Q \begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} = (i) \begin{bmatrix} 1 \\ -i \end{bmatrix} \text{ and } Q \begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = (-i) \begin{bmatrix} 1 \\ i \end{bmatrix}$$

Here is some warnings about eigenvalues and eigenvectors.

- The eigenvalues of $A + B$ are not usually $\lambda(A) + \lambda(B)$
- The eigenvalues of AB are not usually $\lambda(A) \times \lambda(B)$.
- A double eigenvalue $\lambda_1 = \lambda_2$ might or might not have two independent eigenvectors.
- The eigenvectors of a real matrix A are orthogonal if and only if $A^T A = AA^T$.

The matrix A also controls a system of linear differential equations $d\mathbf{u}/dt = A\mathbf{u}$. The system starts at an initial vector $\mathbf{u}(0)$ when $t = 0$.

$$\mathbf{u}(0) = c_1\mathbf{x}_1 + \cdots + c_n\mathbf{x}_n$$
$$\mathbf{u}(t) = c_1 e^{\lambda_1 t}\mathbf{x}_1 + \cdots + c_n e^{\lambda_n t}\mathbf{x}_n$$

Example

$$\begin{cases} y_1' = -0.02y_1 + 0.02y_2 & y_1(0) = 0 \\ y_2' = 0.02y_1 - 0.02y_2 & y_2(0) = 150 \end{cases}$$

Example

$$\begin{cases} y_1' = -0.02y_1 + 0.02y_2 & y_1(0) = 0 \\ y_2' = 0.02y_1 - 0.02y_2 & y_2(0) = 150 \end{cases}$$

$$\Rightarrow \mathbf{y}' = A\mathbf{y} \text{ where } A = \begin{bmatrix} -0.02 & 0.02 \\ 0.02 & -0.02 \end{bmatrix}$$

Example

$$\begin{cases} y_1' = -0.02y_1 + 0.02y_2 & y_1(0) = 0 \\ y_2' = 0.02y_1 - 0.02y_2 & y_2(0) = 150 \end{cases}$$

$\Rightarrow \mathbf{y}' = A\mathbf{y}$ where $A = \begin{bmatrix} -0.02 & 0.02 \\ 0.02 & -0.02 \end{bmatrix}$

Let $\mathbf{y} = \mathbf{x}e^{\lambda t}$, then $\mathbf{y}' = \lambda \mathbf{x}e^{\lambda t} = A\mathbf{x}e^{\lambda t} \Rightarrow A\mathbf{x} = \lambda \mathbf{x} \Rightarrow (A - \lambda I)\mathbf{x} = 0$.

Example

$$\begin{cases} y_1' = -0.02y_1 + 0.02y_2 & y_1(0) = 0 \\ y_2' = 0.02y_1 - 0.02y_2 & y_2(0) = 150 \end{cases}$$

$\Rightarrow \mathbf{y}' = A\mathbf{y}$ where $A = \begin{bmatrix} -0.02 & 0.02 \\ 0.02 & -0.02 \end{bmatrix}$

Let $\mathbf{y} = \mathbf{x}e^{\lambda t}$, then $\mathbf{y}' = \lambda \mathbf{x}e^{\lambda t} = A\mathbf{x}e^{\lambda t} \Rightarrow A\mathbf{x} = \lambda \mathbf{x} \Rightarrow (A - \lambda I)\mathbf{x} = 0$.

$\Rightarrow \lambda_1 = 0, \lambda_2 = -0.04, \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Example

$$\begin{cases} y_1' = -0.02y_1 + 0.02y_2 & y_1(0) = 0 \\ y_2' = 0.02y_1 - 0.02y_2 & y_2(0) = 150 \end{cases}$$

$\Rightarrow \mathbf{y}' = A\mathbf{y}$ where $A = \begin{bmatrix} -0.02 & 0.02 \\ 0.02 & -0.02 \end{bmatrix}$

Let $\mathbf{y} = \mathbf{x}e^{\lambda t}$, then $\mathbf{y}' = \lambda \mathbf{x}e^{\lambda t} = A\mathbf{x}e^{\lambda t} \Rightarrow A\mathbf{x} = \lambda\mathbf{x} \Rightarrow (A - \lambda I)\mathbf{x} = 0$.

$\Rightarrow \lambda_1 = 0, \lambda_2 = -0.04$, $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$\Rightarrow \mathbf{y} = c_1\mathbf{x}_1 e^{\lambda_1 t} + c_2\mathbf{x}_2 e^{\lambda_2 t} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 e^{-0.04t}$

Example

$$\begin{cases} y_1' = -0.02y_1 + 0.02y_2 & y_1(0) = 0 \\ y_2' = 0.02y_1 - 0.02y_2 & y_2(0) = 150 \end{cases}$$

$$\Rightarrow \mathbf{y}' = A\mathbf{y} \text{ where } A = \begin{bmatrix} -0.02 & 0.02 \\ 0.02 & -0.02 \end{bmatrix}$$

Let $\mathbf{y} = \mathbf{x}e^{\lambda t}$, then $\mathbf{y}' = \lambda \mathbf{x} e^{\lambda t} = A\mathbf{x}e^{\lambda t} \Rightarrow A\mathbf{x} = \lambda \mathbf{x} \Rightarrow (A - \lambda I)\mathbf{x} = 0$.

$$\Rightarrow \lambda_1 = 0, \lambda_2 = -0.04, \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\Rightarrow \mathbf{y} = c_1\mathbf{x}_1 e^{\lambda_1 t} + c_2\mathbf{x}_2 e^{\lambda_2 t} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 e^{-0.04t}$$

$$\Rightarrow \mathbf{y}(0) = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 150 \end{bmatrix} \Rightarrow c_1 = 75, c_2 = -75$$

$$\therefore \mathbf{y} = 75 \begin{bmatrix} 1 \\ 1 \end{bmatrix} - 75 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-0.04t}$$

Computing the Eigenvalues (by hand)

It is easy to see that $A\mathbf{x} = \lambda\mathbf{x}$ is equivalent to $(A - \lambda I)\mathbf{x} = 0$.
Then $(A - \lambda I)$ is not **invertible(singular)**.($\det(A - \lambda I) = 0$)

Question

If A is shifted to $A + sI$, what happens to the $\mathbf{x}$'s and λ's?

Answer

The eigenvectors $\mathbf{x}$ stay the same. Every eigenvalue λ shifts by the number s :

$$(A + sI)\mathbf{x} = \lambda\mathbf{x} + s\mathbf{x} = (\lambda + s)\mathbf{x}$$

Similar Matrices

Definition (Similar Matrix)

The matrices BAB^{-1} (for every invertible B) are “similar” to A: same eigenvalues.

Eigenvector of BAB^{-1} associated with λ is the eigenvectors $\mathbf{x}$ of A are multiplied by B.

$$\because \ (BAB^{-1})(B\mathbf{x}) = BA\mathbf{x} = B\lambda\mathbf{x} = \lambda(B\mathbf{x})$$

When the determinant of $A - \lambda I$ would be completely hopeless, it is very hard to compute eigenvalues of large matrices.

We use **triangularization**.

The idea is to make BAB^{-1} gradually into a **triangular matrix**.

The eigenvalues are not changing and they gradually show up on the main diagonal of BAB^{-1}.

When the determinant of $A - \lambda I$ would be completely hopeless, it is very hard to compute eigenvalues of large matrices.
We use **triangularization**.
The idea is to make BAB^{-1} gradually into a **triangular matrix**.
The eigenvalues are not changing and they gradually show up on the main diagonal of BAB^{-1}.

Limitation of triangularization

We have to compute B^{-1} for each invertible matrix B!

Diagonalizing a Matrix

Suppose A has a full set of n independent eigenvectors.
Put those eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ into an invertible matrix X. Then

$$A \begin{bmatrix} \mathbf{x}_1 \cdots \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} A\mathbf{x}_1 \cdots A\mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \lambda\mathbf{x}_1 \cdots \lambda\mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 \cdots \mathbf{x}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

Let Λ be eigenvalue matrix. The equation $AX = X\Lambda$ tells us that $A = X\Lambda AX^{-1}$.

Diagonalizing a Matrix

Suppose A has a full set of n independent eigenvectors.
Put those eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ into an invertible matrix X. Then

$$A \begin{bmatrix} \mathbf{x}_1 \cdots \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} A\mathbf{x}_1 \cdots A\mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \lambda\mathbf{x}_1 \cdots \lambda\mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 \cdots \mathbf{x}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

Let Λ be eigenvalue matrix. The equation $AX = X\Lambda$ tells us that $A = X\Lambda AX^{-1}$.

Limitation of Diagonalization

We still have to compute X^{-1}!

Schur triangularization theorem

What if...

What if B is a unitary matrix in triangularization $T = BAB^{-1}$?
Then we can triangulate the matrix A without finding the inverse matrix B^{-1}. Just take B^*.

Schur Decomposition

If A is a $n \times n$ square matrix with complex entries, then A can be expressed as

$$A = UTU*$$

where U is a unitary matrix, and T is an upper triangular matrix.

Schur triangularization theorem

Definition (unitarily similar)

Two square matrices A and B are *unitarily similar* if there exist unitary matrix P such that

$$A = P^* BP.$$

Theorem

Every square complex matrix A is unitarily similar to an upper triangular matrix, i.e., there exists a unitary matrix U such that $T = U^* AU$ is triangular.

Proof.

We use mathematical induction on size of A.
($n = 1$) trivial.
Assume that $n > 1$, and the result holds for all matrices of size less than n. n. Since every complex matrix has an eigenvalue, choose an eigenvalue λ of A and an associated eigenvector $\mathbf{v} = (v_1, \ldots, v_n)$.
Let $\mathbf{x} = \frac{\overline{v_1}\mathbf{v}}{\|\overline{v_1}\mathbf{v}\|}$ and set $u = \mathbf{x} - e_1$.
And we will put Q in some cases.

$$\begin{cases} Q : \text{Householder matrix associated with } u & (\text{if } \mathbf{x} \neq e_1) \\ Q = I & (\text{if } \mathbf{x} = e_1) \end{cases}$$

Schur triangularization theorem

Proof.

Then $\mathbf{x} = Qe_1$, it means that the first column of Q is $\mathbf{x}$. We already know that every householder matrix is unitary and hermitian. So x^* is first row of Q^*. Since $Q^{-1} = Q^* = Q$, $Q = \left[\mathbf{x}|V\right] = \begin{bmatrix} \mathbf{x}^* \\ V^* \end{bmatrix}$. Therefore,

$$QAQ = QA\left[\mathbf{x}|V\right] = Q\left[\lambda\mathbf{x}|AV\right] = \left[\lambda e_1 \middle| \begin{bmatrix} \mathbf{x}^* \\ V^* \end{bmatrix} AV\right] = \begin{bmatrix} \lambda & \mathbf{x}^* AV \\ \mathbf{0} & V^* AV \end{bmatrix}.$$

Schur triangularization theorem

Proof.

The size of V^*AV is $(n-1)\times(n-1)$, so we can apply the induction, there exists unitary matrix R such that $T_{n-1} = R^*(V^*AV)R$ is upper triangular matrix. Let

$$U = Q\begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & R \end{bmatrix},$$

then

$$U^*U = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & R^* \end{bmatrix} Q^*Q \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & R \end{bmatrix} = I.$$

So U is unitary.

Schur triangularization theorem

Proof.

$$T = U^*AU = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & R^* \end{bmatrix} QAQ \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & R \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & R^* \end{bmatrix} \begin{bmatrix} \lambda & \mathbf{x}^*AV \\ \mathbf{0} & V^*AV \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & R \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & R^* \end{bmatrix} \begin{bmatrix} \lambda & \mathbf{x}^*AVR \\ \mathbf{0} & V^*AVR \end{bmatrix}$$

$$= \begin{bmatrix} \lambda & \mathbf{x}^*AVR \\ \mathbf{0} & R^*V^*AVR \end{bmatrix} = \begin{bmatrix} \lambda & \mathbf{x}^*AVR \\ \mathbf{0} & T_{n-1} \end{bmatrix}$$

Hence, T is triangular matrix. □

Nondiagonalizable Matrices

Suppose λ is an eigenvalue of A.

1. **Eigenvectors (geometric)** There are nonzero solutions to $A\mathbf{x} = \lambda\mathbf{x}$.
2. **Eigenvalues (algebraic)** The determinant of $A - \lambda I$ is zero.

And we want to know its **multiplicity**.

1. **(Geometric Multiplicity = GM)** Count the independent eigenvectors for λ. Look at the dimension of the nullspace of $A - \lambda I$.
2. **(Algebraic Multiplicity = AM)** Count the repetitions of λ among the eigenvalues. Look at the roots of $det(A - \lambda I) = 0$.

Note

Always $GM \leq AM$ for each λ.

Nondiagonalizable Matrices

Proof.

Let $\mathbf{x}_1, \dots, \mathbf{x}_r$ be linearly independent eigenvectors associated to $\hat{\lambda}$, so $\hat{\lambda}$ has geometric multiplicity r. Let $\mathbf{x}_{r+1}, \dots, \mathbf{x}_n$ be basis for $\mathbb{R}^n$.
And let S be the matrix which columns $\mathbf{x}_k$.
Consider AS.

$$AS = \begin{bmatrix} | & & | & \\ \hat{\lambda}\mathbf{x}_1 & \cdots & \hat{\lambda}\mathbf{x}_r & \cdots \\ | & & | & \end{bmatrix} \Rightarrow S^{-1}AS = \begin{bmatrix} \hat{\lambda} & & & \\ & \ddots & & B \\ & & \hat{\lambda} & \\ & 0 & & C \end{bmatrix}$$

where $B : r \times n$ matrix, $C : (n-r) \times (n-r)$ matrix.

Nondiagonalizable Matrices

Proof.

$$\begin{aligned}\det(S^{-1}AS - \lambda I) &= \det(S^{-1}AS - S^{-1}(\lambda I)S) \\ &= \det(S^{-1}(A - \lambda I)S) \\ &= \det(S^{-1})\det(A - \lambda I)\det(S) \\ &= \det(A - \lambda I)\end{aligned}$$

Therefore the characteristic polynomial of A and $S^{-1}AS$ are the same. It is easy to see that the characteristic polynomial of $S^{-1}AS$ has a factor of at least $(\hat{\lambda} - \lambda)^r$.($\because$ determinant of block matrices)
$\therefore$ GM $\leq$ AM. □

Next : I.7 Symmetric Positive Definite Matrices