$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}_{m\times n}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}_{n\times 1}}_{x} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}_{m\times 1}}_{b}$$

CX1104: **Linear Algebra for Computing**

Chap. No : **8.1.4**

Lecture : **Eigen and Singular Values**

Topic : **Eigenvalue Decomposition**

Concept : **The Eigenvalue Decomposition**

Instructor: **A/P Chng Eng Siong**

TAs: **Zhang Su, Vishal Choudhari**

Rev: 11th July 2020

Eigendecomposition of a Matrix

- **Eigendecomposition** of a matrix is also known as **spectral decomposition**.
- It factorises a matrix in terms of its:
  - Eigenvalues and Eigenvectors

- Only a diagonalisable matrix can be factorised in this way.
- A **diagonalisable matrix** is a matrix for which algebraic multiplicity (n_i) equals the geometric multiplicity (m_i) for each of its eigenvalues $(m_i = n_i \forall i)$. This also means that its eigenvectors are independent.

Let A be a square $n \times n$ matrix with n linearly independent eigenvectors q_i (where $i = 1,2,\cdots,n$). Then, A can be factorised as:

$$\mathbf{A} = \mathbf{Q\Lambda Q}^{-1}$$

where,

- Q is the square $n \times n$ matrix whose i^{th} column is the eigenvector q_i of A.
- Λ is the diagonal matrix whose diagonal elements are the corresponding eigenvalues, $\Lambda_{ii} = \lambda_i$.

Ref:
https://en.wikipedia.org/wiki/Eigendecomposition_of_a_matrix

Derivation

The eigendecomposition of a matrix A can be derived from the fundamental property of eigenvectors:

Let q_i be an eigenvector corresponding to the eigenvalue λ_i of matrix A:

Hence, $Aq_i = \lambda_i q_i$.

Let $Q = \underbrace{\begin{bmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ q_1 & q_2 & \cdots & q_n \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix}}_{n \text{ linearly independent eigenvectors of } A}$

Ensured by enforcing:

1. $n_i = m_i$
2. **Consequence of proof in the next slide.**

Ref:
https://en.wikipedia.org/wiki/Eigendecomposition_of_a_matrix

Hence,

$$AQ = \begin{bmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ Aq_1 & Aq_2 & \cdots & Aq_n \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ \lambda_1 q_1 & \lambda_2 q_2 & \cdots & \lambda_n q_n \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ q_1 & q_2 & \cdots & q_n \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}}_{\Lambda}$$

Please do verify!

$$= Q\Lambda$$

We have,

$$AQ = Q\Lambda$$

$$\Longrightarrow \quad A = Q\Lambda Q^{-1}$$ if Q is invertible or all q_i are linearly independent.

Ref:
http://mlwiki.org/index.php/Eigendecomposition

Theorem: distinct eigenvalues means linearly independent eigenvectors

Theorem *If $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p$ be eigenvectors of a matrix A corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_p$, respectively, then $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p$ are linearly independent.*

Proof. Let k be the smallest positive integer such that $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ are linearly independent. If $k = p$, nothing is to be proved. If $k < p$, then $\boldsymbol{v}_{k+1}$ is a linear combination of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$; that is, there exist constants $c_1, c_2, \ldots, c_k$ such that

$$\boldsymbol{v}_{k+1} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_k\boldsymbol{v}_k.$$

Applying the matrix A to both sides, we have

$$\begin{aligned} A\boldsymbol{v}_{k+1} &= \lambda_{k+1}\boldsymbol{v}_{k+1} \\ &= \lambda_{k+1}(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_k\boldsymbol{v}_k) \\ &= c_1\lambda_{k+1}\boldsymbol{v}_1 + c_2\lambda_{k+1}v_2 + \cdots + c_k\lambda_{k+1}\boldsymbol{v}_k; \end{aligned}$$

$$\begin{aligned} Av_{k+1} &= A(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_k\boldsymbol{v}_k) \\ &= c_1A\boldsymbol{v}_1 + c_2A\boldsymbol{v}_2 + \cdots + c_kA\boldsymbol{v}_k \\ &= c_1\lambda_1\boldsymbol{v}_1 + c_2\lambda_2\boldsymbol{v}_2 + \cdots + c_k\lambda_k\boldsymbol{v}_k. \end{aligned}$$

Thus

$$c_1(\lambda_{k+1} - \lambda_1)\boldsymbol{v}_1 + c_2(\lambda_{k+1} - \lambda_2)\boldsymbol{v}_2 + \cdots + c_k(\lambda_{k+1} - \lambda_k)\boldsymbol{v}_k = \mathbf{0}.$$

Since $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ are linearly independent, we have

$$c_1(\lambda_{k+1} - \lambda_1) = c_2(\lambda_{k+1} - \lambda_2) = \cdots = c_k(\lambda_{k+1} - \lambda_k) = 0.$$

Note that the eigenvalues are distinct. Hence

$$c_1 = c_2 = \cdots = c_k = 0,$$

which implies that $\boldsymbol{v}_{k+1}$ is the zero vector $\mathbf{0}$. This is contradictory to that $\boldsymbol{v}_{k+1} \neq \mathbf{0}$. $\square$

Ref:
https://www.math.ust.hk/~mabfchen/Math111/Week11-12.pdf

Example

Take the following 2×2 matrix for example:

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 3 \end{bmatrix}$$

The characteristic polynomial $p(\lambda)$ is:

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

$$A - \lambda I = \begin{bmatrix} 1 - \lambda & 0 \\ 1 & 3 - \lambda \end{bmatrix}$$

Hence, $p(\lambda) = det(A - \lambda I) = (1 - \lambda)(3 - \lambda)$.

The eigenvalues are the roots of the equation:

$$p(\lambda) = 0$$

Hence, the eigenvalues of A are:

$$\lambda_1 = 1,\ \lambda_2 = 3$$

Note that the algebraic multiplicities of each of the eigenvalues is 1, i.e.,

$$n_1 = 1$$
$$n_2 = 1$$

Ref:
https://en.wikipedia.org/wiki/Eigendecomposition_of_a_matrix

Now, let's solve for eigenvectors corresponding to $\lambda_1 = 1$:

We look for those vectors x such that:

$$Ax = \lambda_1 x$$
$$(A - \lambda_1 I)x = 0$$

These vectors are equivalent to those in the null space of $(A - \lambda_1 I)$.

Defining $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$,

$$(A - \lambda_1 I)x = 0$$

$$\begin{bmatrix} 1 - 1 & 0 - 0 \\ 1 - 0 & 3 - 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$\begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

We have:

$$x_1 + 2x_2 = 0$$
$$x_1 = -2x_2$$

Hence, $x = \alpha \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, where α can be any scalar. As the eigenspace corresponding to $\lambda_1 = 1$ is spanned by a single vector, it has a dimension of 1. Hence, geometric multiplicity $m_1 = 1$.

Example

Now, let's solve for eigenvectors corresponding to $\lambda_2 = 3$:

We look for those vectors x such that:

$$Ax = \lambda_2 x$$

$$(A - \lambda_2 I)x = 0$$

These vectors are equivalent to those in the null space of $(A - \lambda_1 I)$.

Defining $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$,

$$(A - \lambda_2 I)x = 0$$

$$\begin{bmatrix} 1-3 & 0-0 \\ 1-0 & 3-3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$\begin{bmatrix} -2 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

We have:

$$x_1 = 0$$

$$x_2 = \text{any } \beta \in R$$

Hence, $x = \beta \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, where β can be any scalar. As the eigenspace corresponding to $\lambda_2 = 3$ is spanned by a single vector, it has a dimension of 1. Hence, geometric multiplicity $m_2 = 1$.

Ref:
https://en.wikipedia.org/wiki/Eigendecomposition_of_a_matrix

i	λ_i	n_i	m_i
1	1	1	1
2	3	1	1

Note: $m_i = n_i \ \forall \ i \in \{1, 2\}$

Hence, the matrix A is diagonalisable.

A can be diagonalised as:

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 3 \end{bmatrix} = \underbrace{\begin{bmatrix} \overset{q_1}{\boxed{\begin{matrix} -2 \\ 1 \end{matrix}}} & \overset{q_2}{\boxed{\begin{matrix} 0 \\ 1 \end{matrix}}} \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} -2 & 0 \\ 1 & 3 \end{bmatrix}^{-1}}_{Q^{-1}}$$

Do verify on MATLAB or Python!

Verification

i	λ_i	n_i	m_i
1	1	1	1
2	3	1	1

Note: $m_i = n_i \; \forall \; i \in \{1, 2\}$

Hence, the matrix A is diagonalisable.

A can be diagonalised as:

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 3 \end{bmatrix} = \underbrace{\begin{array}{cc} q_1 & q_2 \\ \left[\begin{array}{cc} \boxed{\begin{matrix} -2 \\ 1 \end{matrix}} & \boxed{\begin{matrix} 0 \\ 1 \end{matrix}} \end{array}\right] \end{array}}_{Q} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} -2 & 0 \\ 1 & 3 \end{bmatrix}^{-1}}_{Q^{-1}}$$

Do verify on MATLAB or Python!

Ref:
https://en.wikipedia.org/wiki/Eigendecomposition_of_a_matrix

Note: q_i is the eigenvector corresponding to λ_i

```
>> Q = [-2 0; 1 1]

Q =
      q1    q2
      -2     0
       1     1
```

Matrix Q

```
>> L = [1 0; 0 3]

L =

λ1→  1     0
     0     3 ←λ2
```

Matrix Λ

```
>> Q*L*inv(Q)

ans =

     1     0
     1     3   A
```

Matrix $A = Q\Lambda Q^{-1}$

Do verify on MATLAB or Python!

Interpretation of Eigen Decomposition of Matrix A

Case 1: **Transformation Matrix**

The matrix of a linear transformation

The matrix of a linear transformation is a matrix for which $T(\vec{x}) = A\vec{x}$, for a vector $\vec{x}$ in the domain of T. This means that applying the transformation T to a vector is the same as multiplying by this matrix. Such a matrix can be found for any linear transformation T from R^n to R^m, for fixed value of n and m, and is unique to the transformation. In this lesson, we will focus on how exactly to find that matrix A, called the *standard matrix* for the transformation.

Ref:
https://www.mathbootcamps.com/matrix-linear-transformation/

> *Decomposing a matrix in terms of its eigenvalues and its eigenvectors gives valuable insights into the properties of the matrix. Certain matrix calculations, like computing the power of the matrix, become much easier when we use the eigendecomposition of the matrix.*
>
> — Page 262, No Bullshit Guide To Linear Algebra, 2017

Ref:
https://machinelearningmastery.com/introduction-to-eigendecomposition-eigenvalues-and-eigenvectors/

More on this in Lecture 8.2 and 8.3.

Case 2: **Covariance Matrix**

Ref:
https://www.visiondummy.com/2014/04/geometric-interpretation-covariance-matrix/

- The largest eigenvector of the covariance matrix always points into the direction of the largest variance of the data, and the magnitude of this vector equals the corresponding eigenvalue.
- The second largest eigenvector is always orthogonal to the largest eigenvector, and points into the direction of the second largest spread of the data.

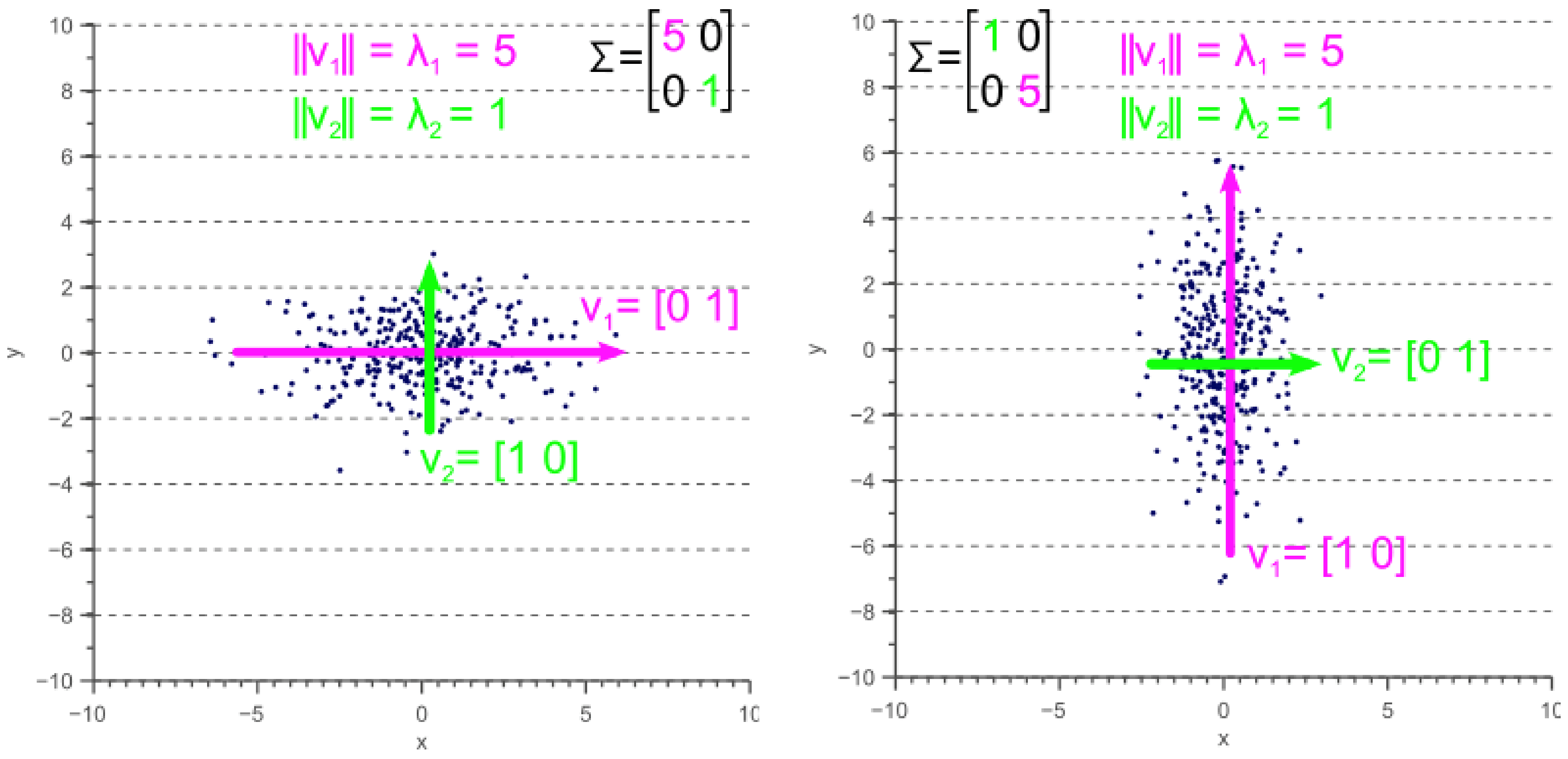

Figure 4. Eigenvectors of a covariance matrix