Matrix Decomposition

✓ LU

✓ QR

→ Spectral (square matrices)

SVD

Eigensystems - Formulation (Part 2)

$$\underline{A}\,\underline{x} = \lambda \underline{x}$$ Eigensystem of $\underline{A}$

λ : Eigenvalue

$\underline{x}$: Eigenvector

$\det(\underline{A} - \lambda \underline{I}) \Rightarrow$ Characteristic polynomial of matrix $\underline{A}$ (in terms of λ)

Properties of Eigensystems:

1. Eigenvalues of $\underline{A}^2$ are the square of the eigenvalues of $\underline{A}$, but eigenvectors are exactly the same

Markov Chains:

$$\underline{P}_n = \underline{M}^n \underline{P}_0$$

may need to raise $\underline{M}$ to a large power; find eigensystem of $\underline{M}$

② Row reduction does not preserve eigenvalues

Row reduction involves the scaling + addition of the matrix rows

③ The product of the eigenvalues of $\underline{A}$ equals det$(\underline{A})$ and the sum of the eigenvalues equals tr$(\underline{A})$, where

$$\text{tr}(\underline{A}) = \text{trace of } \underline{A} = \text{sum of the diagonal}$$

④ One can have imaginary (or complex) eigenvalues, even if $\underline{A}$ is real

Example:

$$\underline{A} = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 2 & 2 \\ 4 & -1 & 0 \end{bmatrix} \Rightarrow \begin{vmatrix} 6-\lambda & 0 & 0 \\ 0 & 2-\lambda & 2 \\ 4 & -1 & -\lambda \end{vmatrix} = 0$$

$$= (6-\lambda)\left[(2-\lambda)(-\lambda) - (-1)(2)\right] = 0$$

$$= (6-\lambda)(\lambda^2 - 2\lambda + 2) = 0$$

$$\lambda = 6 \qquad \lambda = \frac{2 \pm (4-8)^{1/2}}{2}$$

$$\lambda = 1 \pm i$$

$$\therefore \lambda = 6,\ 1+i,\ 1-i$$

Eigenvalues can be complex!

What are the corresponding eigenvectors?

Let $\lambda = 1+i$

$$(\underline{A} - \lambda \underline{I})\underline{x} = \underline{0}$$

$$\begin{bmatrix} 6-(1+i) & 0 & 0 \\ 0 & 2-(1+i) & 2 \\ 4 & -1 & -(i+1) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

rref: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1+i \\ 0 & 0 & 0 \end{bmatrix}$ rank = 2

$\det(\underline{A} - \lambda \underline{I}) = 0$

$\therefore x_1 = 0$

Let $x_2 = 1$, $x_3 = -\frac{x_2}{1+i} \cdot \frac{1-i}{1-i} = -\frac{1-i}{2}$

$$\underline{x} = \begin{bmatrix} 0 \\ 1 \\ -\frac{1}{2}(1-i) \end{bmatrix} \text{ for } \lambda = 1+i$$

Let $\lambda = 1-i$

$$\begin{bmatrix} 6-(1-i) & 0 & 0 \\ 0 & 2-(1-i) & 2 \\ 4 & -1 & -(i-1) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

rref: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1-i \\ 0 & 0 & 0 \end{bmatrix}$ rank : 2

$\det(\underline{A} - \lambda \underline{I}) = 0$

$\therefore x_1 = 0$, $x_2 = 1$, $x_3 = -\frac{1}{1-i} \cdot \frac{1+i}{1+i} = \frac{-1-i}{2}$

$$\underline{x} = \begin{bmatrix} 0 \\ 1 \\ -\frac{1}{2}(1+i) \end{bmatrix} \quad \text{for } \lambda = 1 - i$$

(5) For a real matrix, all complex eigenvalues & eigenvectors will come in complex conjugate pairs

If $\lambda = a + ib$ and $\underline{A}\,\underline{x} = \lambda \underline{x}$ gives the eigenvector $\underline{x}$, then

$$\left(\overline{\underline{A}\,\underline{x}}\right) = \left(\overline{\lambda x}\right) \Rightarrow \overline{\underline{A}}\,\overline{\underline{x}} = \overline{\lambda}\,\overline{\underline{x}}$$

for $\underline{A}$ real $\rightarrow$ $\underline{A}\,\overline{\underline{x}} = \overline{\lambda}\,\overline{\underline{x}}$

$\therefore$ λ has eigenvector $\underline{x}$

$\overline{\lambda}$ has eigenvector $\overline{\underline{x}}$

(6) Repeated eigenvalues are possible

Example:

$$\underline{A} = \begin{bmatrix} 3 & -1 & 2 \\ 3 & -1 & 6 \\ -2 & 2 & -2 \end{bmatrix} \Rightarrow \lambda = -4, \underline{\underline{2, 2}}$$

Then $\lambda = 2$ has a multiplicity of 2

(more precisely, this is the algebraic multiplicity)

Let $\lambda = 2$, Then

$$\underline{A} - \lambda \underline{I} = \begin{bmatrix} 1 & -1 & 2 \\ 3 & -3 & 6 \\ -2 & 2 & -4 \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$\Rightarrow$ 2 free variables

$\Rightarrow \dim(N(\underline{A} - \lambda \underline{I})) = 2$

$\Rightarrow$ 2 eigenvectors

If the # of eigenvectors (or geometric multiplicity) equals the algebraic multiplicity, then that eigenvalue is said to be complete

If the # of eigenvectors is less than the (algebraic) multiplicity, then that eigenvalue is said to be defective

Example:

$$\underline{A} = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix} \rightarrow \lambda = 3, 3$$

Algebraic multiplicity equals 2

$$\underline{A} - \lambda \underline{I} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$ only has

eigenvector $\underline{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ Geometric multiplicity equals 1

$\Rightarrow \lambda = 3$ is defective

If any eigenvalue of $\underline{A}$ is defective, then $\underline{A}$ is said to be defective

---

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

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

Choose $x_1 = 1$

$\begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} x_2 \end{bmatrix} \quad \begin{bmatrix} 0 \end{bmatrix}$

Choose $x_1 = 1$

for $\lambda = 3$, $\quad \underline{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

Normal Matrix Properties

A Normal Matrix is one in which

$$\underline{A}^T \underline{A} = \underline{A}\,\underline{A}^T$$

This occurs for:

→ Symmetric $\underline{A}$ (i.e., $\underline{A} = \underline{A}^T$)

Skew-symmetric $\underline{A}$ (i.e., $\underline{A} = -\underline{A}^T$)

Orthogonal (square) $\underline{A}$ (i.e., $\underline{A}^T \underline{A}$ is a diagonal matrix)

Next, consider real symmetric matrices (but some of the discussion applies to other normal matrices)

Since $\underline{A}$ is real + symmetric, $\underline{A}$ is square

① A real symmetric (normal) matrix, only has real eigenvalues

Proof: Let $\underline{A}$ be real + symmetric, and let λ be any (including possibly complex) eigenvalue, such that $\underline{A}\,\underline{x} = \lambda\,\underline{x}$

$$\underline{A}\,\underline{x} = \lambda\,\underline{x} \quad \& \quad \underline{A}\,\bar{\underline{x}} = \bar{\lambda}\,\bar{\underline{x}}$$

Also, $(\underline{A}\,\bar{\underline{x}})^T = (\bar{\lambda}\,\bar{\underline{x}})^T \Rightarrow \bar{\underline{x}}^T\underline{A}^T = \bar{\underline{x}}^T\bar{\lambda}$

but $\underline{A}^T = \underline{A} \Rightarrow \bar{\underline{x}}^T\underline{A} = \bar{\underline{x}}^T\bar{\lambda}$

Next, take inner product of $\bar{\underline{x}}$ with $\underline{A}\,\underline{x} = \lambda\,\underline{x}$

and inner product of $\underline{x}$ with $\bar{\underline{x}}^T\underline{A} = \bar{\underline{x}}^T\bar{\lambda}$

Thus,

$$\underbrace{\bar{\underline{x}}^T(\underline{A}\,\underline{x})} = \bar{\underline{x}}^T(\lambda\,\underline{x}) \quad \& \quad \underbrace{(\bar{\underline{x}}^T\underline{A})\,\underline{x}} = (\bar{\underline{x}}^T\bar{\lambda})\,\underline{x}$$

equal

so that $\bar{\underline{x}}^T\lambda\,\underline{x} = \bar{\underline{x}}^T\bar{\lambda}\,\underline{x}$

$$\lambda\,\bar{\underline{x}}^T\underline{x} = \bar{\lambda}\,\bar{\underline{x}}^T\underline{x}$$

Since $\bar{\underline{x}}^T\underline{x} = |x_1|^2 + |x_2|^2 + \dots + |x_n^2| > 0$

$\rightarrow \quad \lambda = \bar{\lambda} \quad$ with $\lambda = a + ib$

$\bar{\lambda} = a - ib$

$\therefore$ $b = 0$ and λ must be real

If $\underline{A}$ is real & symmetric, all eigenvalues

λ must be real //

② All eigenvectors of a real symmetric matrix are orthogonal

Proof: Let $\underline{A}\,\underline{x} = \lambda_1 \underline{x}$ & $\underline{A}\,\underline{y} = \lambda_2 \underline{y}$

for $\lambda_1 \neq \lambda_2$

Since $\underline{A}$ is real & symmetric, λ_1 & λ_2 are real. Then

$$(\lambda_1 \underline{x}) \cdot \underline{y} = (\lambda_1 \underline{x})^T \underline{y} = (\underline{A}\,\underline{x})^T \underline{y} = \underline{x}^T \underline{A}^T \underline{y}$$

$$= \underline{x}^T \underline{A}\,\underline{y} = \underline{x}^T (\lambda_2 \underline{y})$$

$$\Rightarrow \underline{x}^T \lambda_1 \underline{y} = \underline{x}^T \lambda_2 \underline{y}$$

$$\lambda_1 (\underline{x}^T \underline{y}) = \lambda_2 (\underline{x}^T \underline{y})$$

Since $\lambda_1 \neq \lambda_2 \Rightarrow$ This is true

iff $\underline{x}^T \underline{y} = 0 \rightarrow$ Eigenvectors are orthogonal //

(3) One can also prove that the eigenvectors for a real, symmetric matrix $\underline{A}$ can be orthonormal

Note: Eigenvectors can be scaled by any non-zero constant, so that constant can be chosen to set the length of each eigenvector to unity

Matrix Diagonalization

Matrix Diagonalization is the application of a matrix $\underline{P}$ such that

Lambda

$$\underline{P}^{-1}\underline{A}\,\underline{P} = \underline{\Lambda}$$, where $\underline{\Lambda}$ is a diagonal matrix

Consider the eigensystem of $\underline{A} \in M_{nn}$

$$\underline{A}\,\underline{x} = \lambda \underline{x}$$

Let $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ represent the eigenvectors of $\underline{A}$ with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$

$$\underline{A}\,\underline{x}_1 = \lambda_1 \underline{x}_1$$

$$\underline{A}\,\underline{x}_2 = \lambda_2 \underline{x}_2$$

$$\vdots$$

$$\underline{A}\,\underline{x}_n = \lambda_n \underline{x}_n$$

Consider

$$\underline{A}\left[\underline{x}_1\ \underline{x}_2\ \cdots\ \underline{x}_n\right] = \underline{A}\,\underline{S} = \left[\lambda_1\underline{x}_1\ \lambda_2\underline{x}_2\ \cdots\ \lambda_n\underline{x}_n\right]$$

$$\underline{A}[\underline{x}_1\ \underline{x}_2 \cdots \underline{x}_n] = \underline{A}\,\underline{S} = [\lambda_1\underline{x}_1\ \lambda_2\underline{x}_2 \cdots \lambda_n\underline{x}_n]$$

Let

$$\underline{\Lambda} = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$ diagonal!

Then

$$[\lambda_1\underline{x}_1\ \lambda_2\underline{x}_2 \cdots \lambda_n\underline{x}_n] = \underline{S}\,\underline{\Lambda}$$

$$\therefore\ \underline{A}\,\underline{S} = \underline{S}\,\underline{\Lambda}$$

a) Let $\underline{A}$ be non-defective $\Rightarrow$ All eigenvalues of $\underline{A}$ are complete

$\Rightarrow$ There are n independent eigenvectors

$\Rightarrow$ $\underline{S}^{-1}$ exists

$\Rightarrow$ $\underline{S}^{-1}\underline{A}\,\underline{S} = \underline{\Lambda}$

or $\underline{A} = \underline{S}\,\underline{\Lambda}\,\underline{S}^{-1}$ $\leftrightarrow$ eigendecomposition of $\underline{A}$

(or spectral decomposition)

b) If $\underline{A}$ is defective, then $\underline{S}^{-1}$ does not exist

$\Rightarrow$ If $\underline{A}$ is defective, one can not find an eigendecomposition

$\Rightarrow$ $\underline{A}$ is not diagonalizable

Let all eigenvalues of $\underline{A}$ be complete (thus $\underline{A}$ is complete) and consider $\underline{A}^2$

$$\underline{A}^2 = \underline{A}\,\underline{A} = (\underline{S}\,\underline{\Lambda}\,\underline{S}^{-1})(\underline{S}\,\underline{\Lambda}\,\underline{S}^{-1})$$
$$= \underline{S}\,\underline{\Lambda}^2\,\underline{S}^{-1}$$

with

$$\underline{\Lambda}^2 = \begin{bmatrix} \lambda_1^2 & & & 0 \\ & \lambda_2^2 & & \\ & & \ddots & \\ 0 & & & \lambda_n^2 \end{bmatrix}$$

More generally, $\underline{A}^m = \underline{S}\,\underline{\Lambda}^m\,\underline{S}^{-1}$

Even more generally, $f(\underline{A}) = \underline{S}\,f(\underline{\Lambda})\,\underline{S}^{-1}$

Even more generally, $f(\underset{\sim}{A}) = \underset{\sim}{S}\, f(\underset{\sim}{\Lambda})\, \underset{\sim}{S}^{-1}$

$$e^{\underset{\sim}{A}} = \underset{\sim}{S}\, e^{\underset{\sim}{\Lambda}}\, \underset{\sim}{S}^{-1}$$

$$\begin{bmatrix} e^{\lambda_1} & & & 0 \\ & e^{\lambda_2} & & \\ & & \ddots & \\ 0 & & & e^{\lambda_n} \end{bmatrix}$$

Example: Markov Chains

$$\underline{P}_m = \underline{M}^m \underline{P}_0 \qquad \text{with } \underline{M} \text{ complete}$$

$$= \underline{S}_M \underline{\Lambda}_M^m \underline{S}_M^{-1} \underline{P}_0$$

$$\underline{\Lambda}_M^m = \begin{bmatrix} \lambda_1^m & & & 0 \\ & \lambda_2^m & & \\ & & \ddots & \\ 0 & & & \lambda_n^m \end{bmatrix}$$

If all eigenvalues λ_i of $\underline{M}$ have $|\lambda_i| \leq 1$, then as $m \to \infty$, $\underline{M}$ will go to a steady state version $\underline{M}^\infty$ dominated by the largest eigenvalue (in an absolute value sense)

For the Markov transition matrix

$$\max |\lambda_i| = 1$$

Calling this λ_1, only the first column of $\underline{S}$ (i.e. the corresponding eigenvector $\underline{x}_1$) and the first row of $\underline{S}^{-1}$ (i.e. the vector $\underline{y}_1$) will determine the steady response. Thus

$$\underline{P}_\infty = \left(\underline{x}_1 \lambda_1 \underline{y}_1\right) \underline{P}_0$$

or

$$\underline{P}_\infty = \underline{S} \begin{bmatrix} \lambda_1 & & & 0 \\ & 0 & & \\ & & 0 & \\ & & & \ddots & \\ 0 & & & & 0 \end{bmatrix} \underline{S}^{-1}$$

Example: Assume all eigenvalues of $\underline{A}$ have $|\lambda_i| < 1$

Then,

$$\lim_{m \to \infty} \underline{A}^m \underline{x} = \underline{0}$$

Now, let $\underline{A}$ be Normal (i.e. $\underline{A}\,\underline{A}^T = \underline{A}^T \underline{A}$)

$\Rightarrow$ All eigenvectors are orthonormal

$\Rightarrow$ $\underline{S}$ = unitary matrix, such that

$$\underline{S}^T \underline{S} = \underline{S}\,\underline{S}^T = \underline{I} \Rightarrow \underline{S}^{-1} = \underline{S}^T$$

and this $\underline{S}$ is denoted as $\underline{\Phi}$

If $\underline{A}$ is normal, then

$$\underline{A} = \underline{\Phi}\,\underline{\Lambda}\,\underline{\Phi}^T$$ unitary decomposition

Any unitary decomposition is the summation of rank 1 matrices

Rank 1 matrix has the form $\underline{u}\,\underline{v}^T$ outer product

$$\underline{A} = \underline{\Phi}\,\underline{\Lambda}\,\underline{\Phi}^T = \begin{bmatrix} \underline{x}_1 & \underline{x}_2 & \cdots & \underline{x}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} \underline{x}_1^T \\ \underline{x}_2^T \\ \vdots \\ \underline{x}_n^T \end{bmatrix}$$

$$= \lambda_1 \underline{x}_1 \underline{x}_1^T + \lambda_2 \underline{x}_2 \underline{x}_2^T + \cdots + \lambda_n \underline{x}_n \underline{x}_n^T$$

or, from another perspective,

$\underline{x}_i \underline{x}_i^T$ is the projection onto the eigenspace with a basis given by $\{\underline{x}_i\}$

What if $\underline{A}$ is non-diagonalizable?

(e.g., $\underline{A}$ is defective)

Use Schur's Theorem, which states that every square matrix $\underline{A}$ can be written as

$$\underline{A} = \underline{Q}\,\underline{T}\,\underline{Q}^T,$$ where $\underline{T}$ is an upper-triangular matrix and $\underline{Q}$ is unitary

Here, $\underline{A}$ and $\underline{T}$ are similar matrices having the same rank, trace, determinant and eigenvalues

Then, because $\underline{T}$ is upper triangular, the eigenvalues of $\underline{A}$ appear on the diagonal of $\underline{T}$

Summary

All square matrices: $\underline{A} = \underline{Q}\,\underline{T}\,\underline{Q}^T$

If $\underline{A}$ is complete: $\underline{A} = \underline{S}\,\underline{\Lambda}\,\underline{S}^{-1}$

If $\underline{A}$ is normal: $\underline{A} = \underline{Q}\,\underline{\Lambda}\,\underline{Q}^T \rightarrow \underline{Q}^T\underline{A}\,\underline{Q} = \underline{\Lambda}$

Now consider real, symmetric $\underline{A}$ with only positive eigenvalues $\longrightarrow$ positive definite

One way to establish positive definiteness is to compute all of the eigenvalues

However, this is very expensive $\mathcal{O}(n^3)$

Another approach: Check $\underline{x}^T\underline{A}\,\underline{x}$

$$\underline{A}\,\underline{x} = \lambda\,\underline{x}$$

$$\underline{x}^T \underline{A}\,\underline{x} = \lambda\,\underline{x}^T \underline{x}$$

with $\underline{x}^T \underline{x} = |x_1|^2 + |x_2|^2 + \ldots + |x_n|^2 > 0$

Therefore, with $\underline{x}^T \underline{A}\,\underline{x} > 0$, we must have $\lambda > 0$

Beyond this, it can be shown that if $\underline{x}^T \underline{A}\,\underline{x} > 0$ for any $\underline{x}$, then all the λ's will be positive

Here, $\underline{x}^T \underline{A}\,\underline{x}$ is the "energy" definition of positive definite