1 How to compute the matrix exponential... and more!

1.1 Distinct eigenvalues

Theorem: If matrix $A \in \mathbb{R}^{n\times n}$ (or $\in \mathbb{C}^{n\times n}$) has m distinct eigenvalues ($\lambda_i \neq \lambda_j$, $\forall i \neq j = 1, \dots, m$) then it has (at least) m linearly independent eigenvectors.

Corollary: If all eigenvalues of A are distinct then A is diagonalizable!

Proof of the Theorem: (By contradiction) Assume λ_i, $i = 1, \dots, m$ are distinct and v_i, $i = 1, \dots, m$ are linearly dependent. That is, there exists α_i such that

$$\sum_{i=1}^{m} \alpha_i v_i = 0$$

where α_i are not zero all zero. We can assume w.l.o.g that $\alpha_1 \neq 0$. Multiplying on the left by $(\lambda_m I - A)$

$$\begin{aligned} 0 = (\lambda_m I - A)\sum_{i=1}^{m} \alpha_i v_i &= (\lambda_m I - A)\sum_{i=1}^{m-1} \alpha_i v_i + \alpha_m(\lambda_m I - A)v_m, \\ &= \sum_{i=1}^{m-1} \alpha_i(\lambda_m - \lambda_i)v_i, \end{aligned}$$

since $Av_i = \lambda_i v_i$. Then multiply by $(\lambda_{m-1} I - A)$

$$(\lambda_{m-1} I - A)\sum_{i=1}^{m-1} \alpha_i(\lambda_m - \lambda_i)v_i = \sum_{i=1}^{m-2} \alpha_i(\lambda_{m-1} - \lambda_i)(\lambda_m - \lambda_i)v_i = 0$$

Repeatedly multiply by $(\lambda_{m-k} I - A)$, $k = 2, \dots, m-2$ to obtain

$$\alpha_1 \prod_{i=2}^{m} (\lambda_i - \lambda_1)v_i = 0.$$

As $\lambda_1 \neq \lambda_i, \forall i = 2, \dots, m$, the above implies $\alpha_1 = 0$, which is a contradiction.

1.2 Non-diagonalizable matrices

Question: Are there matrices which are not diagonalizable by a similarity transformation?

Short answer: YES.

Long answer:

$\star$ If $A = T\Lambda T^{-1}$ then Λ must contain only eigenvalues. Recall that

$$\begin{bmatrix} Av_1 & \cdots & Av_n \end{bmatrix} = AT = T\Lambda = \begin{bmatrix} \lambda_1 v_1 & \cdots & \lambda_n v_n \end{bmatrix}.$$

Matrix $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, has $\lambda_1 = \lambda_2 = 0$ (see homework), therefore $\Lambda = 0$. But if A is diagonalizable then there exists T nonsingular such $T^{-1}\Lambda T = 0 \neq A$!

1.3 Now what?

First, a matrix might have repeated eigenvalues and still be diagonalizable.

Simple example: $A = I$.

Not so simple example: $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

Eigenvalues

$$0 = \det(\lambda I - A) = \begin{vmatrix} \lambda - 1 & 0 & -1 \\ 0 & \lambda - 1 & 0 \\ 0 & 0 & \lambda - 2 \end{vmatrix} = (\lambda - 1)(\lambda - 1)(\lambda - 2)$$

$$\Rightarrow \quad \lambda_1 = \lambda_2 = 1, \quad \lambda_3 = 2.$$

Eigenvectors

$$(1I - A)v = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} v = 0, \qquad \Rightarrow \qquad v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$

$$(2I - A)v = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} v = 0, \qquad \Rightarrow \qquad v_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Therefore

$$\underbrace{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1}}_{T^{-1}} \underbrace{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{T} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}}_{\Lambda}$$

Conclusion: eigenvalues with multiplicity greater than one might have linearly independent eigenvectors.

1.4 Jordan Form

Definition: A *Jordan Block* $J_k(\lambda) \in \mathbb{C}^{k \times k}$ is the upper triangular matrix

$$J_k(\lambda) = \begin{bmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ 0 & & & \lambda \end{bmatrix}.$$

Theorem: For any matrix $A \in \mathbb{R}^{n \times n}$ (or $\in \mathbb{C}^{n \times n}$) there exists T nonsingular such that $J = T^{-1}AT$, where

$$J = \begin{bmatrix} J_{k_1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{k_m}(\lambda_m) \end{bmatrix},$$

the eigenvalues $\lambda_1, \ldots, \lambda_m$ with multiplicity $k_1, \ldots, k_m$ are not necessarily distinct, and $k_1 + \cdots + k_m = n$.

Corollary: When $m = n$ then $J = \Lambda$.

Proof of the Theorem: Let $p \leq m$ be the number of eigenvalues of A with any multiplicity but associated with linearly independent eigenvectors.
For $i = 1, \ldots, p$, let λ_i, be one of such eigenvalues and v_i its associated eigenvector. Set $k_i = 1$ so that $J_{k_i}(\lambda_i) = J_1(\lambda_i) = \lambda_i$, $i = 1, \ldots, p$.

For $i > p$ define $T_i = \begin{bmatrix} v_{i1} & \cdots & v_{ik_i} \end{bmatrix}$ and consider

$$\begin{bmatrix} Av_{i1} & \cdots & Av_{ik_i} \end{bmatrix} = AT_i = T_i J_{k_i}(\lambda_i) = \begin{bmatrix} v_{i1} & \cdots & v_{ik_i} \end{bmatrix} \begin{bmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \lambda_i & 1 \\ 0 & & & \lambda_i \end{bmatrix},$$

which is equivalent to

$$\begin{aligned} Av_{i1} &= \lambda_i v_{i1}, \\ Av_{i2} &= v_{i1} + \lambda_i v_{i2}, \\ \vdots \ &\ \ \vdots \\ Av_{ik_i} &= v_{ik_{i-1}} + \lambda_i v_{ik_i}. \end{aligned}$$

For $k = 2$

$$(A - \lambda_i I) v_{i2} = v_{i1} \neq 0 \quad \Rightarrow \quad v_{i2} \neq 0,$$

which multiplied by $(A - \lambda_i I)$ on the left produces

$$(A - \lambda_i I)^2 v_{i2} = (A - \lambda_i I) v_{i1} = 0.$$

In general

$$\begin{aligned} (A - \lambda_i I)^{j-1} v_{ij} &= v_{ij-1} \neq 0, \quad \Rightarrow \quad v_{ij} \neq 0 \\ (A - \lambda_i I)^j v_{ij} &= 0. \end{aligned}$$

Using the above facts, we can prove that v_{ij}, $j = 1, \ldots, k_i$, are linearly independent, the same way we did for distinct eigenvalues.

Putting it all together: matrix $T = \begin{bmatrix} v_1 & \cdots & v_p & T_{p+1} \cdots & T_m \end{bmatrix}$ is nonsingular and

$$\begin{aligned} AT &= \begin{bmatrix} Av_1 & \cdots & Av_p & AT_{p+1} & \cdots & AT_m \end{bmatrix} \\ &= \begin{bmatrix} v_1 J_1(\lambda_1) & \cdots & v_p J_1(\lambda_p) & T_{p+1} J_{k_{p+1}}(\lambda_{p+1}) & \cdots & T_m J_{k_m}(\lambda_m) \end{bmatrix} = TJ \end{aligned}$$

so that

$$J = T^{-1} A T.$$

1.5 Back to matrix functions

Let $f(x) : \mathbb{C} \to \mathbb{C}$ be a continuous scalar-valued function.

Definition: For any $X \in \mathbb{R}^{n\times n}$ (or $\in \mathbb{C}^{n\times n}$)

$$f(X) = T \begin{bmatrix} f(J_{k_1}(\lambda_1)) & & 0 \\ & \ddots & \\ 0 & & f(J_{k_m}(\lambda_m)) \end{bmatrix} T^{-1},$$

where T nonsingular and $J_{k_1}, \cdots, J_{k_m}$ are such that

$$\begin{bmatrix} J_{k_1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{k_m}(\lambda_m) \end{bmatrix} = T^{-1} A T.$$

1.6 What is $f(J_k(\lambda))$?

The simplest particular case is $J_2(\lambda)$. Define

$$J_2(\lambda, \epsilon) = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda + \epsilon \end{bmatrix},$$

with eigenvalues

$$\lambda_1 = \lambda, \qquad \lambda_2 = \lambda + \epsilon.$$

For any $\epsilon \neq 0$, $J_2(\lambda, \epsilon)$ is diagonalizable.
Computing eigenvectors

$$[\lambda_1 I - J_2(\lambda, \epsilon)]v_1 = \begin{bmatrix} 0 & -1 \\ 0 & -\epsilon \end{bmatrix} v_1 = 0 \quad \Rightarrow \quad v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$
$$[\lambda_2 I - J_2(\lambda, \epsilon)]v_2 = \begin{bmatrix} \epsilon & -1 \\ 0 & 0 \end{bmatrix} v_2 = 0, \quad \Rightarrow \quad v_2 = \begin{pmatrix} 1 \\ \epsilon \end{pmatrix}.$$

and

$$T = \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & \epsilon \end{bmatrix}, \qquad T^{-1} = \begin{bmatrix} 1 & -1/\epsilon \\ 0 & 1/\epsilon \end{bmatrix}$$

we can evaluate

$$\begin{aligned} f(J_2(\lambda, \epsilon)) = T f(\Lambda) T^{-1} &= \begin{bmatrix} 1 & 1 \\ 0 & \epsilon \end{bmatrix} \begin{bmatrix} f(\lambda) & 0 \\ 0 & f(\lambda + \epsilon) \end{bmatrix} \begin{bmatrix} 1 & -1/\epsilon \\ 0 & 1/\epsilon \end{bmatrix}, \\ &= \begin{bmatrix} f(\lambda) & (f(\lambda + \epsilon) - f(\lambda))/\epsilon \\ 0 & f(\lambda + \epsilon) \end{bmatrix}. \end{aligned}$$

As $J_2(\lambda, \epsilon) \to J_2(\lambda)$ as $\epsilon \to 0$ and f is continuous, if f is also differentiable at λ

$$f(J_2(\lambda, \epsilon)) = \lim_{\epsilon \to 0} f(J_2(\lambda, \epsilon)) = \begin{bmatrix} f(\lambda) & f'(\lambda) \\ 0 & f(\lambda) \end{bmatrix}.$$

In the general case

$$f(J_k(\lambda)) = \begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2!} f''(\lambda) & \cdots & \frac{1}{(k-1)!} f^{(k-1)}(\lambda) \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \frac{1}{2!} f''(\lambda) \\ & & & \ddots & f'(\lambda) \\ 0 & & & & f(\lambda) \end{bmatrix}.$$

1.7 What is $f(J_k(\lambda)t)$?

Again for $k = 2$

$$
\begin{aligned}
f(J_2(\lambda,\epsilon)t) = T f(\Lambda t) T^{-1} &= \begin{bmatrix} 1 & 1 \\ 0 & \epsilon \end{bmatrix} \begin{bmatrix} f(\lambda t) & 0 \\ 0 & f((\lambda+\epsilon)t) \end{bmatrix} \begin{bmatrix} 1 & -1/\epsilon \\ 0 & 1/\epsilon \end{bmatrix}, \\
&= \begin{bmatrix} f(\lambda t) & (f((\lambda+\epsilon)t) - f(\lambda t))/\epsilon \\ 0 & f((\lambda+\epsilon)t) \end{bmatrix}.
\end{aligned}
$$

And before we take the limit let $\delta = \epsilon t$ so that

$$
\lim_{\epsilon \to 0} f(J_2(\lambda,\epsilon)t) = \lim_{\delta \to 0} \begin{bmatrix} f(\lambda t) & t(f(\lambda t + \delta) - f(\lambda t))/\delta \\ 0 & f(\lambda t + \delta) \end{bmatrix} = \begin{bmatrix} f(\lambda t) & t f'_\lambda(\lambda t) \\ 0 & f(\lambda t) \end{bmatrix},
$$

In the general case

$$
f(J_k(\lambda t)) = \begin{bmatrix} f(\lambda t) & t f'(\lambda t) & \frac{t^2}{2!} f''(\lambda t) & \cdots & \frac{t^{k-1}}{(k-1)!} f^{(k-1)}(\lambda t) \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \frac{t^2}{2!} f''(\lambda t) \\ & & & \ddots & t f'(\lambda t) \\ 0 & & & & f(\lambda t) \end{bmatrix}.
$$

Example

$$
e^{J_k(\lambda t)} = \begin{bmatrix} e^{\lambda t} & t e^{\lambda t} & \frac{t^2}{2!} e^{\lambda t} & \cdots & \frac{t^{k-1}}{(k-1)!} e^{\lambda t} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \frac{t^2}{2!} e^{\lambda t} \\ & & & \ddots & t e^{\lambda t} \\ 0 & & & & e^{\lambda t} \end{bmatrix}.
$$

1.8 Example:

Compute e^{At} for $A = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix}$.

Eigenvalues:

$$\det(\lambda I - A) = \begin{vmatrix} \lambda & -1 \\ 1 & \lambda + 2 \end{vmatrix} = \lambda^2 + 2\lambda + 1 = (\lambda + 1)^2 = 0$$

$$\Rightarrow \qquad \lambda_1 = \lambda_2 = -1.$$

Eigenvector:

$$(1I - A)v_1 = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} v_1 = 0, \quad \Rightarrow \quad v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Generalized eigenvector:

$$v_{11} = v_1, \;\; (A - 1I)v_{12} = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} v_{12} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = v_{11}, \quad \Rightarrow \quad v_{12} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Jordan form:

$$T = \begin{bmatrix} v_{11} & v_{12} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}, \qquad T^{-1} = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}, \qquad J = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}.$$

Matrix exponential:

$$\begin{aligned} e^{At} &= Te^{J_2(-t)}T^{-1}, \\ &= \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} e^{-t} & te^{-t} \\ 0 & e^{-t} \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}, \\ &= \begin{bmatrix} (t+1)e^{-t} & te^{-t} \\ -te^{-t} & (1-t)e^{-t} \end{bmatrix}. \end{aligned}$$