A Review of some Linear Algebra Topics

1 Eigenvalues and Eigenvectors

Let A be a $n \times n$ matrix whose elements are members of the field K, ($K = \mathbb{R}$ or $\mathbb{C}$), $\lambda \in K$ and $\mathbf{e} \neq \mathbf{0}$ a n-vector such that

$$A\mathbf{e} = \lambda\mathbf{e} \tag{1}$$

then λ is an *eigenvalue* of A, and $\mathbf{e}$ a corresponding eigenvector.
From Eq(1)

$$(\lambda I - A)\mathbf{e} = \mathbf{0}$$

and so λ must satisfy the CHARACTERISTIC equation:

$$det(\lambda I - A) = 0 \tag{2}$$

It can be shown that

$$det(\lambda I - A) = \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0 \tag{3}$$

This is called the characteristic polynomial of A.
Note: In theory, Eq(2) can be used to find λ 's – than Eq(1) used to find the corresponding $\mathbf{e}$'s.
Example 1:

$$A = \begin{pmatrix} 0 & 1 \\ -2 & -3 \end{pmatrix}$$

$$\begin{aligned} det(\lambda I - A) &= det \begin{pmatrix} \lambda & -1 \\ 2 & \lambda + 3 \end{pmatrix} \\ &= \lambda^2 + 3\lambda + 2 \\ &= 0 \Rightarrow \lambda = -1 \text{ or } -2 \end{aligned}$$

For $\lambda_1 = -1$

$$\begin{pmatrix} 0 & 1 \\ -2 & -3 \end{pmatrix} \mathbf{e}_1 = (-1)\mathbf{e}_1 \Rightarrow \mathbf{e}_1 = \alpha_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \alpha_1 \in K$$

For $\lambda_2 = -2$

$$\begin{pmatrix} 0 & 1 \\ -2 & -3 \end{pmatrix} \mathbf{e}_2 = (-2)\mathbf{e}_2 \Rightarrow \mathbf{e}_2 = \alpha_2 \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \quad \alpha_2 \in K$$

2 Similarity

Two $n \times n$ matrices R and S are said to be *similar* if there exist an *invertible* matrix P such that

$$R = P^{-1}SP$$

Similar matrices have the same *spectrum* (set of eigenvalues):

$$\begin{aligned} det(\lambda I - R) &= det(\lambda I - P^{-1}SP) \\ &= det(\lambda P^P - P^{-1}SP) \\ &= det\left(P^{-1}(\lambda I - S)P\right) \\ &= detP^{-1}det(\lambda I - S)detP \\ &= det(\lambda I - S) \end{aligned}$$

3 Diagonalisability

Recall that a diagonal matrix is a $n \times n$ matrix, all of whose off-diagonal entries are zero. We denote a diagonal matrix by $diag\{d_1, d_2, \ldots, d_n\}$ where $d_1, d_2, \ldots, d_n$ are the diagonal entries.

The $n \times n$ matrix A is said to be *diagonalisable* if it is similar to a diagonal matrix. It can be shown that the diagonal entries of the diagonal matrix are the eigenvalues of A.

Not every square matrix is diagonalisable. A necessary and sufficient condition for A to be diagonalisable is that its eigenvectors form a *linearly independent* set.
Let A have spectrum $\lambda_1, \lambda_2, \ldots, \lambda_n$ with associated eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ respectively. Then, for $i = 1, 2, \ldots, n$ we have $A\mathbf{e}_i = \lambda_i \mathbf{e}_i$. Consider

$$\begin{aligned} A\,E = A\,[\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n] &= [A\mathbf{e}_1, A\mathbf{e}_2, \ldots, A\mathbf{e}_n] \\ &= [\lambda_1\mathbf{e}_1, \lambda_2\mathbf{e}_2, \ldots, \lambda_n\mathbf{e}_n] \\ &= [\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n]\, diag\{\lambda_1, \lambda_2, \ldots, \lambda_n\} \\ &= E\,\Lambda \end{aligned}$$

where $E = [\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n]$ and $\Lambda = diag\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$.
Thus, if A is diagonalisable

$$A = E\,\Lambda\,E^{-1} \tag{4}$$

Λ is unique up to ordering of the eigenvalues.

We note that it can be shown that eigenvectors corresponding to distinct eigenvalues are linearly independent; hence, if A has n distinct eigenvalues, it is diagonalisable.
Example 1 (Cont'd): $A = E\,\Lambda\,E^{-1}$

$$\begin{pmatrix} 0 & 1 \\ -2 & -3 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix}$$

We also note that the following corollary of Eq(4) which may be proven using induction

$$A^k = E\,\Lambda^k\,E^{-1} \qquad k = 0, 1, \ldots \tag{5}$$

where it is also straightforward to show that

$$\Lambda^k = diag\{\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k\} \tag{6}$$

4 Jordan Canonical Form

Although every square matrix is not diagonalisable, it is possible to show that every matrix A is similar to a Jordan Form matrix J, i.e.

$$A = P^{-1}\,J\,P$$

where J is a block diagonal matrix

$$J = diag\{J_1, J_2, \ldots, J_s\} = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & J_s \end{pmatrix}$$

with each block J_i being of size $n_i \times n_i$ with $\sum n_i = n$ and of form

$$J_i = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}$$

where λ belongs to the spectrum of A. (The same eigenvalue may appear in more than one block of J). The Jordan Form of A is unique up to the ordering of the blocks. If A is diagonalisable, then its Jordan Form coincides with its diagonalised form.

It is straightforward to establish that

$$A^k = P^{-1}\, J^k\, P \tag{7}$$

where $J^k = diag\{J_1^k, J_2^k, \ldots, J_s^k\}$ and

$$J_i^k = \begin{pmatrix} \lambda^k & c_k(1)\lambda^{k-1} & c_k(2)\lambda^{k-2} & \cdots & c_k(n_i-1)\lambda^{k-n_i-1} \\ 0 & \lambda^k & c_k(1)\lambda^{k-1} & \cdots & c_k(n_i-2)\lambda^{k-n_1-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda^k \end{pmatrix} \tag{8}$$

where $c_k(j) = \binom{k}{j}$. Example 2 :

$$\hat{A} = \begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix}$$

has eigenvalue $\lambda = -1$ (multiplicity 2) and associated eigenvector $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Hence it is not diagonalisable since there are not two linear independent eigenvectors. However ($\hat{A} = P^{-1}\, J\, P$)

$$\begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 1/2 \\ -1 & 1/2 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1/2 & -1/2 \\ 1 & 1 \end{pmatrix}$$

5 Cayley-Hamilton Theorem

"Every square matrix satisfies its own characteristic equation".
Example 1 (cont'd): $A = \left(\begin{smallmatrix} 0 & 1 \\ -2 & -3 \end{smallmatrix}\right)$ has characteristic polynomial $\chi(\lambda) = \lambda^2 + 3\lambda + 2$. Hence the theorem says

$$\begin{aligned} A^2 + 3A + 2I &= 0 \\ \Rightarrow \begin{pmatrix} -2 & -3 \\ 6 & 7 \end{pmatrix} + \begin{pmatrix} 0 & 3 \\ -6 & -9 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Example 2 (cont'd): $\hat{A} = \left(\begin{smallmatrix} 0 & 1 \\ -1 & -2 \end{smallmatrix}\right)$ has characteristic polynomial $\chi(\lambda) = \lambda^2 + 2\lambda + 1$. Here the theorem says

$$\begin{aligned} \hat{A}^2 + 2\hat{A} + I &= 0 \\ \Rightarrow \begin{pmatrix} -1 & -2 \\ 2 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ -2 & -4 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

A corollary of the theorem is then: There exist scalars $s_0(k), s_1(k), \ldots, s_{n-1}(k)$ such that

$$A^k = \sum_{j=0}^{n-1} s_j(k) A^j$$

for all $k \geq 0$.

6 Computing e^{At}

By definition

$$e^{At} = \sum_{k=0}^{\infty} A^k \frac{t^k}{k!}$$

From Eq(7) this gives

$$e^{At} = P^{-1} \left(\sum_{k=0}^{\infty} diag\{J_1^k, J_2^k, \ldots, J_s^k\} \frac{t^k}{k!} \right) P = P^{-1}\, diag\{e^{J_1}, e^{J_2}, \ldots, e^{J_s}\}\, P$$

where it is straightforward but tedious to establish (using Eq(8)) that

$$e^{J_i} = \sum_{k=0}^{\infty} J_i^k \frac{t^k}{k!} = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2}e^{\lambda t} & \cdots & \frac{t^{n_i-1}}{(n_i-1)!}e^{\lambda t} \\ 0 & e^{\lambda t} & te^{\lambda t} & \cdots & \frac{t^{n_i-2}}{(n_i-2)!}e^{\lambda t} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & e^{\lambda t} \end{pmatrix}$$

When A is diagonalisable, Eq(5) gives

$$e^{At} = E \left(\sum_{k=0}^{\infty} diag\{\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k\} \frac{t^k}{k!} \right) E^{-1} = E\, diag\{e^{\lambda_1 t}, e^{\lambda_2 t}, \ldots, e^{\lambda_n t}\}\, E^{-1} \tag{9}$$

Example 1 (cont'd): Since A is diagonalisable

$$\begin{aligned} e^{At} &= E\, diag\{e^{\lambda_1 t}, e^{\lambda_2 t}\}\, E^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} e^{-t} & 0 \\ 0 & e^{-2t} \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 2e^{-t} - e^{-2t} & e^{-t} - e^{-2t} \\ -2e^{-t} + 2e^{-2t} & -e^{-t} + 2e^{-2t} \end{pmatrix} \end{aligned}$$

Example 2 (cont'd): Since $\hat{A}$ is not diagonalisable

$$\begin{aligned} e^{\hat{A}t} &= P^{-1}\, e^{Jt}\, P \\ &= \begin{pmatrix} 1 & 1/2 \\ -1 & 1/2 \end{pmatrix} \begin{pmatrix} e^{-t} & te^{-t} \\ 0 & e^{-t} \end{pmatrix} \begin{pmatrix} 1/2 & -1/2 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} (1+t)e^{-t} & te^{-t} \\ -te^{-t} & (1-t)e^{-t} \end{pmatrix} \end{aligned}$$

7 Rank of a Matrix

Let M be a $p \times q$ matrix (of real entries). If we view the rows of the matrix as vectors in the space $\mathbb{R}^q$. the number of linearly independent vectors in this set is called the *row rank* of M. Similarly, if we view the columns of M as vectors in $\mathbb{R}^p$, the number of linearly independent vectors in this set is called the *column rank* of M. A standard result tells us that these two ranks are equal, and so we talk about the *rank* of M.

One method of computing the rank of M is to perform a series of row operations on M (e.g. as in the elimination phase of the *Gauss* Elimination algorithm) which reduces M to *row echelon form.* The number of non-zero rows is then the rank. Of course, this procedure could also be applied to M^T - why?

Example: Consider the 3×4 matrix

$$\begin{pmatrix} 4 & -2 & 5 & 1 \\ 2 & 6 & -3 & -1 \\ 1 & 7 & -2 & -4 \end{pmatrix}$$

Reduction to row echelon form yields after the first pass

$$\begin{pmatrix} 4 & -2 & 5 & 1 \\ 0 & 7 & -11/2 & -3/2 \\ 0 & 15/2 & -13/4 & -17/4 \end{pmatrix}$$

and then after the second pass

$$\begin{pmatrix} 4 & -2 & 5 & 1 \\ 0 & 7 & -11/2 & -3/2 \\ 0 & 0 & 37/14 & -37/14 \end{pmatrix}$$

Thus rank = 3.

The rank of M must satisfy

$$\text{rank} \leq \min(p, q)$$

A matrix is said to be of *full rank* if rank $= \min(p, q)$. When M is square ($p = q$), then M is of full rank (i.e. rank $= p$) if and only if $\det M \neq 0$.

8 When does the linear system $M\mathbf{y} = \mathbf{b}$ have a solution?

Let $M = [\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_q]$ be a $p \times q$ matrix where $\mathbf{m}_j$ is the j-th column of M, let $\mathbf{y} = [y_1, y_2, \ldots, y_q]^T$ the $q \times 1$ the vector of unknowns and $\mathbf{b} = [b_1, b_2, \ldots, b_p]^T$ the $p \times 1$ vector of "right hand sides". We can rewrite the system of equations as

$$\mathbf{b} = \mathbf{m}_1 y_1 + \mathbf{m}_2 y_2 + \cdots + \mathbf{m}_q y_q$$

i.e, $\mathbf{b}$ must be expressible as a linear combination of the columns of M, which leads to the condition that if $\mathbf{b}$ can be any element of a particular subspace, then $\{\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_q\}$ must span this subspace. In particular, if $\mathbf{b}$ may be any element of $\mathbb{R}^p$ then the columns of M must span this space, i.e. M must be be of full rank.