Matrix Method

Contents

1 Matrix Algebra

1.1 Notations and Matrices

If $\boldsymbol{A}$ is an $n \times p$ matrix, it can be presented as:

- $\boldsymbol{A}$: $n \times p$
- $\boldsymbol{A} = \{a_{ij}\}$ where $i \leq n$ and $j \leq p$

Row and column vector representation of $\boldsymbol{A}$:

- row vector as column vector: $\boldsymbol{a}_{(i)}$
- column vector: $\boldsymbol{a}_j$

1.2 Diagonal Matrix

- $\boldsymbol{D}$: $n \times n$ where $d_{ij} = 0$ when $i \neq j$
- $\boldsymbol{D} = diag(d_i)$
- $\boldsymbol{D} = \{d_{ii}\}$
- if $\boldsymbol{A}$ conforms with $\boldsymbol{D}$, then

$$\boldsymbol{D}\boldsymbol{A} = \{d_i a_{ij}\}$$

1.3 Transpose

- if $\boldsymbol{A} = \{a_{ij}\}$, then $\boldsymbol{A}' = \{a_{ji}\}$
- $\boldsymbol{A}$: $n \times p$ and $\boldsymbol{A}'$: $p \times n$
- $(\boldsymbol{A}\boldsymbol{B})' = \boldsymbol{B}'\boldsymbol{A}'$

1.4 Upper and Lower Triangular Matrix

- $\boldsymbol{A} : n \times n$ where $a_{ij} = 0$ for $i > j$ (upper)
- $\boldsymbol{A} : n \times n$ where $a_{ij} = 0$ for $i < j$ (lower)

1.5 Symmetric Matrix

- $\boldsymbol{S}$: $n \times n$ where $\{s_{ij}\} = \{s_{ji}\}$
- $\boldsymbol{S} = \boldsymbol{S}'$
- if $\boldsymbol{A}$ and $\boldsymbol{B}$ are both symmetric, then $(\boldsymbol{AB})' = \boldsymbol{A}'\boldsymbol{B}' = \boldsymbol{BA}$
- for any matrix $\boldsymbol{S}$, $\boldsymbol{S}'\boldsymbol{S}$ and $\boldsymbol{SS}'$ are symmetric matrices

1.6 Identity Matrix

- $\boldsymbol{I}$: $n \times n$
- $\boldsymbol{I} = diag(\mathbf{1})$

1.7 Idempodent Matrix

- $\boldsymbol{A}$: $n \times n$
- $\boldsymbol{A}^2 = \boldsymbol{A}$

1.8 Trace

- if $\boldsymbol{A}$: $n \times n$, then
$$tr(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii},\ tr(\boldsymbol{AA}') = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2$$
- $tr(\boldsymbol{AB}) = tr(\boldsymbol{BA})$ if $\boldsymbol{B}$ is also a square matrix
- $tr(\boldsymbol{A}+\boldsymbol{B}) = tr(\boldsymbol{A}) + tr(\boldsymbol{B})$
- $tr(c\boldsymbol{A}) = c \cdot tr(\boldsymbol{A})$
- $tr(\boldsymbol{AB}') = tr(\boldsymbol{A}'\boldsymbol{B}) = tr(\boldsymbol{B}'\boldsymbol{A}) = tr(\boldsymbol{BA}') = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ij}$

1.9 General Notations

- $\boldsymbol{A} + \boldsymbol{B} = \{a_{ij} + b_{ij}\}$
- $\boldsymbol{cA} = \{c \cdot a_{ij}\}$

1.10 Matrix Multiplication

Given $\boldsymbol{A}$: $n \times p$ and $\boldsymbol{B}$: $r \times c$

$$\begin{aligned} \boldsymbol{C} = \boldsymbol{AB} &= \boldsymbol{A}[\boldsymbol{b}_1\ \boldsymbol{b}_2\ ...\ \boldsymbol{b}_c] \\ &= [\boldsymbol{Ab_1}\ \boldsymbol{Ab_2}\ ...\ \boldsymbol{Ab_c}] \end{aligned}$$

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

1.11 Other Multiplications

- Hadamard product (element-wise): $\boldsymbol{A} \circ \boldsymbol{B}$
- Kronecker product (direct): $\boldsymbol{A} \otimes \boldsymbol{B}$
- Cross product: $\boldsymbol{A} \times \boldsymbol{B}$
- Dot product: $\boldsymbol{A} \cdot \boldsymbol{B}$

1.12 2×2 Determinant

- $\boldsymbol{A}$: 2×2
- $det(\boldsymbol{A}) = |\boldsymbol{A}| = a_{11}a_{22} - a_{12}a_{21}$
- given $\boldsymbol{x},\ \boldsymbol{y}$: 2×1 and $\boldsymbol{A}$: 2×2 and that $\boldsymbol{y} = \boldsymbol{Ax}$, $\boldsymbol{A}$ is said to transform coordinates in an $\boldsymbol{x}{-}space$ to coordinates in a $\boldsymbol{y}{-}space$
- if $a_{11} > a_{21} > 0$ and $a_{22} > a_{12} > 0$, then the unit square in an $\boldsymbol{x}{-}space(x,\ y \geq 0)$ gets transformed into a **parallelogram**
- area of the parallelogram is equal to $|\boldsymbol{A}| > 0$

- geometrically, the determinant is the factor by which the area is scaled in the linear transformation described by the matrix
- when $\boldsymbol{A} < 0$, the sign indicates a directional shift

1.13 Cofactor Matrix

- $\boldsymbol{A} : n \times n$
- $\boldsymbol{C} = (-1)^{i+j}|\boldsymbol{M}_{(ij)}|$, where the $\boldsymbol{M} : (n-1) \times (n-1)$ is obtained from $\boldsymbol{A}$ by deleting row i and column j

1.14 n×n Determinant

- $\boldsymbol{A}$: $n \times n$
- for $n > 2$, we use the recursive formula

$$|\boldsymbol{A}| = \sum_{j}^{n} a_{ij}(-1)^{i+j}|\boldsymbol{M}_{(ij)}|$$

  to determine $|\boldsymbol{A}|$ for any arbitrary $i \leq n$
- $|\boldsymbol{A}| = |\boldsymbol{A'}|$
- $|\boldsymbol{AB}| = |\boldsymbol{BA}| = |\boldsymbol{A}||\boldsymbol{B}|$ iff $\boldsymbol{B}$: $n \times n$
- $|\boldsymbol{A^{-1}}||\boldsymbol{A}| = 1$
- the addition of a multiple of one row (or column) of $\boldsymbol{A}$ to another row (or column) leaves the determinant unchanged (forming an upper triangular or lower triangular matrix)
- swapping columns/rows changes sign of determinant
- multiplying a row/column with a constant increases the determinant by that same factor
- if $\boldsymbol{A}$ is orthogonal, then $|\boldsymbol{A}|^2 = |\boldsymbol{A}||\boldsymbol{A'}| = |\boldsymbol{AA'}| = 1$ such that $|\boldsymbol{A}| \pm 1$

1.15 Invertible Matrix

- given $\boldsymbol{A} : n \times n$, if there exists ($\boldsymbol{B}$: $n \times n$) such that

$$\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$$

  then $\boldsymbol{B} = \boldsymbol{A^{-1}}$ is said to be the **inverse** of $\boldsymbol{A}$

- for any set of linear equations, if the inverse exists, then:

$$\begin{aligned} \boldsymbol{Ax} &= \boldsymbol{b} \\ \boldsymbol{A^{-1}Ax} &= \boldsymbol{b} \\ \boldsymbol{x} &= \boldsymbol{A^{-1}b} \end{aligned}$$

- $(\boldsymbol{BA})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A^{-1}}$

- $(\boldsymbol{A^{-1}})' = (\boldsymbol{A'})^{-1}$

- suppose we can find a sequence of elementary operator matrices $\boldsymbol{E}_{(i)}$, $i \leq K$, such that

$$\boldsymbol{E}_{(K)}\boldsymbol{E}_{(K-1)} \ ... \ \boldsymbol{E}_{(2)}\boldsymbol{E}_{(1)}\boldsymbol{A} = \boldsymbol{I}$$

  then, by definition,

$$\boldsymbol{E}_{(K)}\boldsymbol{E}_{(K-1)} \ ... \ \boldsymbol{E}_{(2)}\boldsymbol{E}_{(1)} = \boldsymbol{A^{-1}}$$

- $\boldsymbol{A^{-1}}$ exists iff $\boldsymbol{A} \neq 0$ (non-singular)

1.16 Adjugate Matrix

- $\boldsymbol{A} : n \times n$

- $adj(\boldsymbol{A}) = \boldsymbol{C'}$

- $\boldsymbol{A^{-1}} = \frac{1}{|\boldsymbol{A}|}\boldsymbol{C'}$

1.17 Singular Matrix

- consider $\boldsymbol{A}$: $n \times n$
- if $det(\boldsymbol{A}) = 0$, then $\boldsymbol{A}$ is said to be **singular** and **not invertible**

1.18 Bilinear and Quadratic Forms

- given a symmetric matrix $\boldsymbol{A} : n \times n$, $\boldsymbol{x} : n \times 1$, and $\boldsymbol{y} : n \times 1$:

$$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{y} = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i a_{ij} y_j$$

  is a **bilinear form**

- if $\boldsymbol{x} = \boldsymbol{y}$ and $\boldsymbol{A}$ is symmetric, then it is the **quadratic form**

1.19 Definiteness

- $\boldsymbol{A}$: $n \times n$ is **positive definite** if $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} > 0$ for any $\boldsymbol{x} \neq \boldsymbol{0}$
- $\boldsymbol{A}$: $n \times n$ is **positive semi-definite** if $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \geq 0$ for at least one $\boldsymbol{x} = \boldsymbol{0}$
- $\boldsymbol{A}$: $n \times n$ is **negative definite** if $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} < 0$ for any $\boldsymbol{x} \neq \boldsymbol{0}$
- $\boldsymbol{A}$: $n \times n$ is **negative semi-definite** if $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \leq 0$ for at least one $\boldsymbol{x} = \boldsymbol{0}$
- $\boldsymbol{A}$: $n \times n$ is **indefinite** if it is not **positive semi-definite** or **negative semi-definite**

1.20 Centring Matrix

- define $\boldsymbol{1}$ as a vector of 1s and $\boldsymbol{J} = \{1_{ij}\}$
- define $\boldsymbol{x}$: $n \times 1$
- $(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J})$ is the centring matrix
- Properties: symmetric, positive semi-definite, singular, idempodent

1.21 Orthogonal Vector

- given $\boldsymbol{x}$: $n \times 1$ and $\boldsymbol{y}$: $n \times 1$, $\boldsymbol{x}$ and $\boldsymbol{y}$ are **orthogonal vectors** if $\boldsymbol{x}'\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{x} = 0$

1.22 Norm or Length

- norm or length of $\boldsymbol{x}$ is given by

$$||\boldsymbol{x}|| = \sqrt{\boldsymbol{x}'\boldsymbol{x}}$$

- unit vector, $\boldsymbol{u} = \frac{\boldsymbol{x}}{||\boldsymbol{x}||}$, has norm/length of 1 and is said to be **normalised**

1.23 Orthonormal Vectors

- if $\boldsymbol{x}$ and $\boldsymbol{y}$ are orthogonal and normalised, they are an **orthonormal pair of vectors**. If the columns of the matrix $\boldsymbol{P}$ are an orthonormal set of vectors, then $\boldsymbol{P}'\boldsymbol{P} = \boldsymbol{I}$. Similarly, if the rows form an orthonormal set then $\boldsymbol{P}\boldsymbol{P}' = \boldsymbol{I}$

1.24 Orthogonal Matrix

- if $\boldsymbol{A}$: $n \times n$ and the rows and columns are mutually orthonormal, $\boldsymbol{A}$ is an **orthogonal matrix**
- $\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}' = \boldsymbol{I}$
- $\boldsymbol{A}' = \boldsymbol{A}^{-1}$
- $|\boldsymbol{A}| = \pm 1$

1.25 Elementary Operations

- **elementary operators** are matrices obtained from making one alteration to $\boldsymbol{I}$
- consider $\boldsymbol{P}_{(ij;\lambda)}$, identical to $\boldsymbol{I}$, except that $p_{ij} = \lambda$ when $i \neq j$
- adds multiple, λ, of i-th column to the one of j-th column
- consider $\boldsymbol{E}_{(ij)}$, identical to $\boldsymbol{I}$, except that the i-th column/row and j-th column/row are swapped

- swaps i-th column/row with j-th column/row
- consider $\boldsymbol{R}_{(i;\lambda)}$, identical to $\boldsymbol{I}$, except that the i-th diagonal element is λ
- multiplies the i-th row/column by λ
- $|\boldsymbol{P}_{(ij;\lambda)}| = 1$
- $|\boldsymbol{E}_{(ij)}| = -1$
- $|\boldsymbol{R}_{(i;\lambda)}| = \lambda$

2 Linear Algebra

2.1 Linear Independence

- consider $\boldsymbol{X}$: $m \times n$ (a vector of vectors) and $\boldsymbol{a}$: $n \times 1$ (a vector of constants)
- $\boldsymbol{X} = [\boldsymbol{x}_1 \ \boldsymbol{x}_2 \ ... \ \boldsymbol{x}_n]$
- $\boldsymbol{X}\boldsymbol{a} = \sum_{i=1}^{n} a_i \boldsymbol{x}_i = \boldsymbol{0}$
- $\boldsymbol{a} = \boldsymbol{0}$ is the trivial solution
- if $\boldsymbol{a} = \boldsymbol{0}$ is the only solution, then the vectors $\boldsymbol{x}_i$ are **linearly independent**
- if at least one $a_i \neq 0$, then the vectors $\boldsymbol{x}_i$ are **linearly dependent**
- rearranging $\sum_{i=1}^{n} a_i \boldsymbol{x}_i = 0$:

$$\boldsymbol{x}_n = -\frac{a_1}{a_n}\boldsymbol{x}_1 - ... - \frac{a_{n-1}}{a_n}\boldsymbol{x}_{n-1}$$

  such that $a_n \neq 0$ (at least one coefficient being non-zero)

- repeat process until r **linearly independent** vectors have been found
- $\boldsymbol{X}$ can be partitioned as:

$$\boldsymbol{X} = [\boldsymbol{X}_1 \ \boldsymbol{X}_2]$$

where $\boldsymbol{X}_1$: $m \times (n-r)$ is a matrix holding $(n-r)$ **linearly dependent** vectors and $\boldsymbol{X}_2$: $m \times r$ is a matrix holding r **linearly independent**

- for some matrix $\boldsymbol{B}$: $r \times (n-r)$, we can write:

$$\boldsymbol{X}_1 = \boldsymbol{X}_2\boldsymbol{B}$$

2.2 Gram-Schmidt Process

2.3 Linearly Independent Rows/Columns

- consider $\boldsymbol{X}$: $m \times n$ (a vector of vectors) and $\boldsymbol{a}$: $n \times 1$ (a vector of constants)
- suppose also that:

$$\boldsymbol{x}_n = a_1\boldsymbol{x}_1 + ... + a_{n-1}\boldsymbol{x}_{n-1}$$

- it is possible to add multiples of the other columns to any one chosen column of $\boldsymbol{X}$ in such a way that this column vector becomes $\mathbf{0}$ such that:

$$\boldsymbol{X}^* = [\mathbf{0}\ \boldsymbol{x}_2\ ...\ \boldsymbol{x}_n]$$

- if columns of $\boldsymbol{X}$ are **linearly dependent**, then $|\boldsymbol{X}| = 0$
- $|\boldsymbol{X}| = |\boldsymbol{X}^*| = 0$. This means:
  - $\boldsymbol{X}^{\mathbf{-1}}$ does not exist (invertible and singular)
  - the linear equation $\boldsymbol{X}\boldsymbol{a} = \boldsymbol{b}$ does not have a valid solution $\boldsymbol{a}$
  - there is some $\boldsymbol{a} \neq \mathbf{0}$ such that $\boldsymbol{X}\boldsymbol{a} = \mathbf{0}$
  - for square matrics, either there is a non-trivial solution to $\boldsymbol{X}\boldsymbol{a} = \mathbf{0}$ or a solution to $\boldsymbol{X}\boldsymbol{a} = \boldsymbol{b} \neq \mathbf{0}$

2.4 Matrix Rank

- consider $\boldsymbol{X}$: $m \times n$

- **rank** of $\boldsymbol{X}$, $rank(\boldsymbol{X})$, is the number of **linearly independent** rows/columns

- $rank(\boldsymbol{X}) \leq min(m,\ n)$

- $\boldsymbol{X}$ is of **full row rank** if $rank(\boldsymbol{X}) = m$ and **full column rank** if $\boldsymbol{X} = n$

- if a square matrix is of full row rank, then it is also of full column rank (and vice versa), and $\boldsymbol{X}$ is said to be of **full rank**. This implies that:

  - $\boldsymbol{X^{-1}}$ does exist (invertible)

  - $|\boldsymbol{X}| \neq 0$ (non-singular)

  - the linear equation $\boldsymbol{Xa} = \boldsymbol{b} \neq \boldsymbol{0}$ has a valid solution $\boldsymbol{a}$

  - $\boldsymbol{Xa} = \boldsymbol{0}$ for $\boldsymbol{a} = \boldsymbol{0}$

- one can find the $rank(\boldsymbol{X})$ by finding $\boldsymbol{X^*}$, the **row echelon form** of $\boldsymbol{X}$, using elementary operations

- $rank(\boldsymbol{X})$ = the number of non-zero rows of $\boldsymbol{X^*}$

2.5 Matrix Factorization and Canonical Forms

- consider $\boldsymbol{A}$: $p \times q$

- let $rank(\boldsymbol{A}) = r$, and suppose that the rows and columns have been ordered such the first r columns and rows of $\boldsymbol{A}$ are linearly independent

- $\boldsymbol{A}$ can be partitioned as follows:

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{X} : r \times r & \boldsymbol{Y} : r \times (q-r) \\ \boldsymbol{Z} : (p-r) \times r & \boldsymbol{W} : (p-r) \times (q-r) \end{bmatrix}$$

  where $rank(\boldsymbol{X}) = r$ (full rank)

- the factorized form of $\boldsymbol{A}$ is:

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{F} \end{bmatrix} \begin{bmatrix} \boldsymbol{X} & \boldsymbol{XH} \end{bmatrix}$$

***where

- $\boldsymbol{A}$ can be transformed by elementary operations into a matrix consisting only of 0's, except that the first r diagonal elements are 1's.

- using the relevant elemtary operations, we can get:

$$\boldsymbol{PAQ} = \begin{bmatrix} \boldsymbol{I} : r \times r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} = \boldsymbol{C}$$

- $\boldsymbol{C}$ is the equivalent **canoncical form** of $\boldsymbol{A}$

- if two matrices reduce to the same canonical form, then they are said to be **equivalent**

- since $\boldsymbol{P}$ and $\boldsymbol{Q}$ are invertible, then:

$$\boldsymbol{A} = \boldsymbol{P^{-1}CQ^{-1}}$$

- if $\boldsymbol{A}$ is not singular, then:

$$\boldsymbol{C} = \boldsymbol{I},\ \boldsymbol{A} = \boldsymbol{P^{-1}Q^{-1}}$$
$$\boldsymbol{A} = \boldsymbol{BB'}$$

where $\boldsymbol{B} = \boldsymbol{P^{-1}}$

- if $\boldsymbol{A}$ is symmetric, it is possible to find a $\boldsymbol{P}$ such that:

$$\boldsymbol{PAP'} = \boldsymbol{C}$$

3 Generalised Inverses and Linear Equations

3.1 Recall

When a column of a matrix in row echelon form contains a pivot, it is called a **basic column**. When it does not contain a pivot, we say that it is a **non-basic column**.

3.2 Full Row Rank

- consider $\boldsymbol{A}$: $m \times n$ of full row rank
- $rank(\boldsymbol{A}) = m$
- since we have m linear equations, so we will aim to solve for m elements of $\boldsymbol{x}$, by arbitrarily setting $(n - m)$ elements to 0.
- $\boldsymbol{A}$ can be partitioned as follows:

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{S} : m \times m & \boldsymbol{T} : m \times (n-m) \end{bmatrix}$$

  where $\boldsymbol{S}$ consists of basic columns and $\boldsymbol{T}$ consists of non-basic columns. We can simplify the linear equation:

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} &= \boldsymbol{b} \\ \begin{bmatrix} \boldsymbol{S} & \boldsymbol{T} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_S \\ \boldsymbol{x}_T \end{bmatrix} &= \boldsymbol{b} \\ \begin{bmatrix} \boldsymbol{S} & \boldsymbol{T} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_S \\ \boldsymbol{0} \end{bmatrix} &= \boldsymbol{b} \\ \boldsymbol{S}\boldsymbol{x}_S + \boldsymbol{T} \cdot \boldsymbol{0} &= \boldsymbol{b} \\ \boldsymbol{x}_S &= \boldsymbol{S}^{-\mathbf{1}}\boldsymbol{b} \end{aligned}$$

- adding back the 0s:

$$\boldsymbol{x} = \boldsymbol{G}\boldsymbol{b}$$
$$= \begin{bmatrix} \boldsymbol{S}^{-1} \\ \\ \boldsymbol{0} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

3.3 Full Column Rank

- consider $\boldsymbol{A}$: $m \times n$ of full column rank
- $rank(\boldsymbol{A}) = n$
- since we have more linear equations, m, than variables, n, to solve, there will be no general solution unless some redundancy exists
- we could find a "closest" approximation to a solution in a least squares sense, by finding the vector $\boldsymbol{x}$ that minimizes:

$$(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x})^{'}(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x})$$

- differentiating with respect to $\boldsymbol{x}$, we get:

$$2\boldsymbol{A}^{'}(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}) = \boldsymbol{0}$$
$$\boldsymbol{x} = (\boldsymbol{A}^{'}\boldsymbol{A})^{-1}\boldsymbol{A}^{'}\boldsymbol{b}$$

- $\boldsymbol{A}^{'}\boldsymbol{A}$ has full rank

3.4 Square Matrix with no Full Rank

- consider $\boldsymbol{A}$: $m \times m$ of full column rank
- $rank(\boldsymbol{A}) < m$
- a solution to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ for $\boldsymbol{b} \neq \boldsymbol{0}$ if $rank(\boldsymbol{A}) = m$

- when $rank(\boldsymbol{A}) < m$, a solution will only exist if the linear relationships between the rows of $\boldsymbol{A}$ are mirrored by equivalent relationships between the elements of $\boldsymbol{b}$
- if $rank([\boldsymbol{A} \;\; \boldsymbol{b}]) = rank(\boldsymbol{A})$, then the aforementioned mirrored relationship is present (non-zero solution exists)

3.5 Generalized Inverse

- consider $\boldsymbol{A}$: $m \times n$
- if $\boldsymbol{AGA} = \boldsymbol{A}$, then $\boldsymbol{AGA}$ is a **generalised inverse** of $\boldsymbol{A}$, $\boldsymbol{A}^-$
- additionally, it is possible that $\boldsymbol{GAG} = \boldsymbol{G}$
- if the two above simultanueously hold, then $\boldsymbol{G}$ is the **reflexive generalized inverse** of $\boldsymbol{A}$
- the three above hold and if $\boldsymbol{AG}$ and $\boldsymbol{GA}$ are symmetric, then $\boldsymbol{G}$ is the **Moore-Penrose inverse** of $\boldsymbol{A}$

3.6 Finding Generalised Inverses

- consider $\boldsymbol{A}$: $m \times n$
- $\boldsymbol{A}$ can be reduced to:

$$\boldsymbol{A}^* = \begin{bmatrix} \boldsymbol{D} : r \times r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}$$

- $\boldsymbol{A}^* = \boldsymbol{PAQ} = \boldsymbol{C}$ (canonical form of $\boldsymbol{A}$)
- the above makes:

$$\boldsymbol{G} = \boldsymbol{Q} \begin{bmatrix} \boldsymbol{D} : r \times r & \boldsymbol{X} \\ \boldsymbol{Y} & \boldsymbol{Z} \end{bmatrix} \boldsymbol{P}$$

  the generalized inverse of $\boldsymbol{A}$ where $\boldsymbol{G}$, $\boldsymbol{Y}$ $\boldsymbol{Z}$ are matrices of arbitrary sizes

- consider

$$\boldsymbol{A} = \boldsymbol{Q} \begin{bmatrix} \boldsymbol{D} : r \times r & \boldsymbol{X} \\ \boldsymbol{Y} & \boldsymbol{Z} \end{bmatrix} \boldsymbol{P}$$

  where $\boldsymbol{D}$ is a non-singular submatrix of $\boldsymbol{A}$

- if $\boldsymbol{Z} = \boldsymbol{Y}\boldsymbol{D}^{-1}\boldsymbol{X}$, then:

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{D}^{-1} : r \times r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}$$

  is the generalized inverse of $\boldsymbol{A}$

3.7 Using Generalized Inverse to Solve Linear Equations

- consider $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$

- if $\boldsymbol{A}$ is not singular, then there exist a single unique solution

- the generalised solution can be expressed as:

$$\boldsymbol{x} = (\boldsymbol{G}\boldsymbol{A} - \boldsymbol{I})\boldsymbol{z}$$

  for a conformable $\boldsymbol{z}$

- if $rank(\boldsymbol{A}) < n$, then this solution, in general, will be non-trivial

- for any arbitrary conformable matrix $\boldsymbol{G}$, we have:

$$(\boldsymbol{A}\boldsymbol{G})\boldsymbol{A}\boldsymbol{x} = (\boldsymbol{A}\boldsymbol{G})\boldsymbol{b}$$

  and if $\boldsymbol{G}$ is a generalised inverse of $\boldsymbol{A}$, then this becomes:

$$\boldsymbol{x} = \boldsymbol{G}\boldsymbol{b}$$

- since $\boldsymbol{x} = (\boldsymbol{G}\boldsymbol{A} - \boldsymbol{I})\boldsymbol{z}$ is a solution to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ for any arbitrary vector $\boldsymbol{z}$, and $\boldsymbol{G}$ is a generalised inverse of $\boldsymbol{A}$, then:

$$\boldsymbol{Gb} + (\boldsymbol{GA} - \boldsymbol{I})\boldsymbol{z}$$

is also a solution

- solution is unique iff $\boldsymbol{GA} = \boldsymbol{I}$

3.8 Eigenvalues

- used to find vectors such that, for some matrix multiplication, the vector changes in such a way that it still lies on the same line
- consider $\boldsymbol{A}$: $n \times n$ and a conformable $\boldsymbol{x} \neq 0$ such that:

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

- if a vector $\boldsymbol{x}$ and scalar λ satisfies this equation, then we refer to them respectively as an **eigenvector** and corresponding **eigenvalue** of $\boldsymbol{A}$
- the scaling of $\boldsymbol{x}$ is arbitrary, since if the above is true, then $\boldsymbol{A}(k\boldsymbol{x}) = \lambda(k\boldsymbol{x})$
- a non-trivial solution will only exist if $(\boldsymbol{A} - \lambda \boldsymbol{I})$ is singular and yielding the **characteristic equation**:

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = 0$$

- it defines an n-th order polynomial in λ
-

3.9 Eigenvectors

- for each of the n eigenvalues of $\boldsymbol{A}$, $\boldsymbol{\lambda} = \{\lambda_1,\ \lambda_2,\ \dots,\ \lambda_n\}$ (set is not unique), there are corresponding eigenvectors $\boldsymbol{U} = \{\boldsymbol{u}_1,\ \boldsymbol{u}_2,\ \dots,\ \boldsymbol{u}_n\}$ (set is unique)

- by constructing $\boldsymbol{U}$: $n \times n$ having the eigenvectors as columns, we have that:

$$\boldsymbol{A}\boldsymbol{U} = \boldsymbol{A}[\boldsymbol{u}_1 \ ... \ \boldsymbol{u}_n] = \boldsymbol{\lambda} \circ \boldsymbol{U} = \boldsymbol{U}\boldsymbol{D}$$

  where $\boldsymbol{D} = diag(\lambda_i)$

- if $\boldsymbol{U}$ is non-singular, then:

$$\boldsymbol{U}^{-\mathbf{1}}\boldsymbol{A}\boldsymbol{U} = \boldsymbol{D}, \ \boldsymbol{A} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^{-\mathbf{1}}$$

- in general:

$$\boldsymbol{A}^k\boldsymbol{x} = \lambda^k\boldsymbol{x}$$

  for some constant k

- if $\boldsymbol{A}$ is non-singular, all the eigenvalues will be non-zero
- we can then say that the pair $(\lambda^{-1}, \ \boldsymbol{x})$ is an **eigenvalue-eigenvector pair** for $\boldsymbol{A}^{-1}$
- $\boldsymbol{A}^k = \boldsymbol{U}\boldsymbol{D}^k\boldsymbol{U}^{-\mathbf{1}}$ and $\boldsymbol{D}^k = diag(\lambda_i^k)$
- $tr(\boldsymbol{A}) = \sum\limits_{i=1}^{n} \lambda_i$
- $|\boldsymbol{A}| = \prod\limits_{i=1}^{n} \lambda_i$
- $rank(\boldsymbol{A}) =$ number of non-zero eigenvalues
- $\lambda = 0$ can be an eigenvalue of $\boldsymbol{A}$, but this implies that $|\boldsymbol{A}| = 0$
- for symmetric matrices, all eigenvalues are real
- suppose we have two pairs $\{\lambda_1, \ \boldsymbol{u}_1\}$ and $\{\lambda_2, \ \boldsymbol{u}_2\}$ of a symmetric matrix $\boldsymbol{A}$ such that $\lambda_1 \neq \lambda_2$. Then:

$$\lambda_1\boldsymbol{u}_2\boldsymbol{u}_2 = \lambda_2\boldsymbol{u}_2\boldsymbol{u}_2$$

- this means that

$$\boldsymbol{u}_2\boldsymbol{u}_2 = \mathbf{0}$$

$\boldsymbol{u}_2$ and $\boldsymbol{u}_1$ are orthogonal vectors

- symmetric matrices with unique eigenvalues have orthogonal eigenvectors
- use Gram-Schmidt process to prove that, even with repeated values of eigenvalues, orthogonal sets of eigenvectors can be found when $\boldsymbol{A}$ is symmetric
- if $\boldsymbol{A}$ is symmetric, then:
  - $\boldsymbol{U}'\boldsymbol{U} = \boldsymbol{I}$
  - $\boldsymbol{U}' = \boldsymbol{U^{-1}}$
  - $\boldsymbol{A} = \boldsymbol{UDU}' = \boldsymbol{UDU^{-1}}$

3.10 Spectral Decomposition

- consider symmetric $\boldsymbol{A}$: $n \times n$
- spectral decomposition of $\boldsymbol{A}$ can be written as:

$$\begin{aligned} \boldsymbol{A} &= \boldsymbol{UDU}' \\ &= \sum_{i=1}^{n} \lambda_i \boldsymbol{u}_i' \boldsymbol{u}_i \end{aligned}$$

- spectral decomposition expresses $\boldsymbol{A}$ as a sum of n rank-1 matrices

3.11 Singular Value Decomposition

- consider $\boldsymbol{X}$: $n \times p$ and $rank(\boldsymbol{X}) \leq p < n$
- $\boldsymbol{X}$ can be expressed in the form:

$$\boldsymbol{X} = \boldsymbol{U}^* \boldsymbol{D}^* \boldsymbol{V}'$$

where $\boldsymbol{U^*}$: $n \times n$, $\boldsymbol{V}$: $p \times p$ are orthogonal matrices and:

$$\boldsymbol{D^*} = \begin{bmatrix} \boldsymbol{D} : k \times k & \boldsymbol{0} : k \times (p-k) \\ \boldsymbol{0} : (n-k) \times k & \boldsymbol{0} : (n-k) \times (p-k) \end{bmatrix}$$

where $\boldsymbol{D} = diag(d_i) = diag(\sqrt{\lambda_i})$

- this is the **full version** of the SVD since it can also be written as the compact SVD:

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}'$$

  where $\boldsymbol{U}$ consists of the first p columns of $\boldsymbol{U^*}$

- since $\boldsymbol{U}$: $n \times p$, its columns are mutually orthonormal. That is:

$$\boldsymbol{U}'\boldsymbol{U} = \boldsymbol{I},\ \boldsymbol{U}\boldsymbol{U}' \neq \boldsymbol{I}$$

- $\boldsymbol{U}$ is the eigenvectors of $\boldsymbol{X}\boldsymbol{X}'$, since $\boldsymbol{X}\boldsymbol{X}'\boldsymbol{U} = \boldsymbol{U}\boldsymbol{D^2}$
- $\boldsymbol{V}$ is the eigenvectors of $\boldsymbol{X}'\boldsymbol{X}$, since $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{V} = \boldsymbol{V}\boldsymbol{D^2}$
- $\boldsymbol{D}$ contains the singular values of $\boldsymbol{X}$ on the diagonal
- $\boldsymbol{D^2}$ contains the eigenvalues of $\boldsymbol{X}\boldsymbol{X}'$ and $\boldsymbol{X}'\boldsymbol{X}$ on the diagonal
- positive quantities $\{\lambda_1,\ \lambda_2,\ \dots\ ,\ \lambda_n\}$ are the non-zero eigenvalues of $\boldsymbol{X}'\boldsymbol{X}$ and the non-zero eigenvalues of $\boldsymbol{X}\boldsymbol{X}'$
- columns of $\boldsymbol{U}$ and $\boldsymbol{V}$ are called the **left and right singular vectors** of $\boldsymbol{X}$ respectively
- $\boldsymbol{D}$ is a diagonal matrix containing the singular values on the diagonal, ordered in deacreasing order
- without loss of generality, the singular values are always positive
- eigenvectors are defined in an arbitrary directional sense: if $\boldsymbol{x}$ is an eigenvector of $\boldsymbol{A}$, then so is $-\boldsymbol{x}$. One needs to check that the directions of the eigenvectors in $\boldsymbol{U}$ and in $\boldsymbol{V}$ are consistently defined