

1 Álgebra Lineal

1.1 Conceptos básicos

Definición 1.1 Una matriz A es un array bidimensional de números reales. Los elementos de A están indexados por sus filas y columnas: a_{ij} es el elemento en la i ésima fila y la j ésima columna. El conjunto de todas las matrices n por m se denota como $\mathbb{R}^{n \times m}$. La transpuesta de una matriz $A \in \mathbb{R}^{n \times m}$ es $A^T \in \mathbb{R}^{m \times n}$ definida como $[a^T]_{ij} = a_{ji}$. Los vectores son un caso especial de matrices con una fila o una columna.

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & \ddots & \ddots & a_{2m} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \dots & \dots & a_{nm} \end{pmatrix} = [a_{ij}]_{n \times m}$$

Definición 1.2 Multiplicando una matriz A por un escalar $c \in \mathbb{R}$ es definido

$$[cA]_{ij} = ca_{ij}$$

La adición de dos matrices A y B del mismo tamaño es definido como

$$[A + B]_{ij} = a_{ij} + b_{ij}.$$

Definición 1.3 Una matriz que tiene el mismo número de filas y columnas se denomina matriz cuadrada. Los elementos diagonales de una matriz cuadrada A son a_{11}, \dots, a_{nn} . Una matriz diagonal es una matriz cuadrada cuyos elementos fuera de la diagonal son cero. Una matriz diagonal cuyos elementos diagonales son todos 1 se llama matriz de identidad y se denota I . Dado un vector de valores diagonales \mathbf{v} denotamos la matriz diagonal correspondiente como **diag** (\mathbf{v}).

Definición 1.4 Una matriz triangular superior es una matriz cuadrada, en el cual todos los elementos por debajo de la diagonal son cero. Una matriz triangular inferior es una matriz cuadrada, en el cual todos los elementos por encima de la diagonal son cero. Una matriz triangular es una matriz triangular inferior o una matriz triangular superior.

Definición 1.5 El producto de una matriz $A \in \mathbb{R}^{n \times m}$ y una matriz $B \in \mathbb{R}^{m \times k}$ es la matriz $AB \in \mathbb{R}^{n \times k}$, definida como

$$[AB]_{ij} = \sum_{r=1}^m a_{ir} b_{rj}.$$

Se cumple las siguientes operaciones

- Asociatividad: $(AB)C = A(BC)$
- Distributividad: $A(B + C) = AB + AC$
- No Conmutatividad: Si $A = [a_{ij}]_{n \times p}$ y $B = [b_{ij}]_{p \times m}$, por definición tenemos $AB = [c_{ij}]_{n \times m}$; sin embargo, BA no está definido si $n \neq m$. En el caso $m = n$, $AB = [c_{ij}]_{n \times n}$ y $BA = [c_{ij}]_{p \times p}$ pero no necesariamente $n = p$, es decir $AB \neq BA$ en general.

Las multiplicaciones repetidas se denotan usando un exponente, por ejemplo $AA = A^2$ y $AA^k = A^{k+1}$.

Ejemplo 1.1 Para $A \in \mathbb{R}^{n \times m}$, $\mathbf{x} \in \mathbb{R}^{m \times 1}$ y $\mathbf{y} \in \mathbb{R}^{1 \times n}$, tenemos

$$\begin{aligned} A\mathbf{x} &\in \mathbb{R}^{n \times 1}, & [A\mathbf{x}]_i &= \sum_{r=1}^m a_{ir}x_r \\ \mathbf{y}A &\in \mathbb{R}^{1 \times m}, & [\mathbf{y}A]_i &= \sum_{r=1}^n y_r a_{ri} \\ \mathbf{y}A\mathbf{x} &\in \mathbb{R}^{1 \times 1}, & \mathbf{y}A\mathbf{x} &= \sum_{i=1}^n \sum_{j=1}^m y_i a_{ij} x_j. \end{aligned}$$

Si $n = m$, tenemos

$$\mathbf{x}^T A \mathbf{x} \in \mathbb{R}^{1 \times 1} \quad \mathbf{x}^T A \mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n x_i A_{ij} x_j.$$

Definición 1.6 Para dos vectores $\mathbf{x} \in \mathbb{R}^{n \times 1}$, $\mathbf{y} \in \mathbb{R}^{m \times 1}$, definimos el producto exterior como

$$\mathbf{x} \otimes \mathbf{y} = \mathbf{xy}^T \in \mathbb{R}^{n \times m},$$

y si $n = m$, definimos el producto interno como

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} = \mathbf{y}^T \mathbf{x} \in \mathbb{R}.$$

Definición 1.7 Nos referimos al vector $\sum_{i=1}^n \alpha_i \mathbf{v}^{(i)}$, donde $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ como una combinación lineal de los vectores $\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(n)}$.

Definición 1.8 Los vectores $\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(n)}$ se dicen que son linealmente independientes, si la expresión $\sum_i \alpha_i \mathbf{v}^{(i)} = 0$ implica que los $\alpha_i = 0$ para todo i . Los vectores que no son linealmente independientes se dicen que son linealmente dependientes. La definición implica que si $\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(n)}$ son linealmente dependientes, entonces $\mathbf{v}^{(i)}$ puede ser expresado como una combinación lineal de los vectores restantes.

Ejemplo 1.2 Los vectores $\mathbf{e}^{(i)} \in \mathbb{R}^{n \times 1}$ definido por $e_j^{(i)} = 1$ si $i = j$ y 0 en otros casos, $i = 1, \dots, n$ son linealmente independientes. Desde que cada vector \mathbf{v} ,

$$\mathbf{v} = \sum v_i \mathbf{e}^{(i)},$$

el conjunto de todas las combinaciones de $\mathbf{e}^{(i)}$, $i = 1, \dots, n$ es \mathbb{R}^n . Los vectores $\mathbf{e}^{(i)}$, $i = 1, \dots, n$ forman una base estándar de \mathbb{R}^n .

Desde la multiplicación matricial, vemos que multiplicar una matriz por un vector de columna $A\mathbf{v}$ produce un vector columna que es una combinación lineal de las columnas de A . Además, la multiplicación matricial AB es una matriz cuyas columnas son cada una, una combinación lineal de los vectores de columna de A .

La siguiente proposición ayuda a explicar por qué las matrices y el álgebra lineal son tan importantes

Proposición 1.1 Cualquier transformación lineal $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ se realiza mediante la multiplicación matricial $T(\mathbf{v}) = A\mathbf{v}$ para algún $A \in \mathbb{R}^{m \times n}$. Por el contrario, la aplicación $\mathbf{v} \mapsto A\mathbf{v}$ es una transformación lineal.

Definición 1.9 La inversa de una matriz cuadrada A es la matriz A^{-1} tal que $AA^{-1} = A^{-1}A = I$. Una matriz A para el cual la inversa existe es llamada no singular. En otro caso es llamada singular.

Proposición 1.2 La inversa de una matriz no singular es única.

Una de las aplicaciones de la inversión de la matriz es resolver sistemas lineales de ecuaciones. Dado el sistema de ecuaciones $A\mathbf{x} = \mathbf{y}$ donde A , \mathbf{y} son conocidos y \mathbf{x} no se conoce, podemos resolver para \mathbf{x} invirtiendo la matriz

$$\mathbf{x} = A^{-1}A\mathbf{x} = A^{-1}\mathbf{y}.$$

Así $A^{-1}\mathbf{y}$ es el vector de coeficientes de la expansión de \mathbf{y} en una combinación lineal usando las columnas de A .

Proposición 1.3 Para dos matrices A, B , tenemos $(AB)^T = B^T A^T$. Similarmente, para matrices no singulares A, B , tenemos $(AB)^{-1} = B^{-1}A^{-1}$.

Proposición 1.4 Para una matriz no singular A , tenemos

$$(A^T)^{-1} = (A^{-1})^T.$$

Definición 1.10 Una matriz A es ortogonal si es una matriz cuadrada, cuyas columnas son vectores ortonormales, o equivalentemente A satisface $AA^T = I$.

Proposición 1.5 La multiplicación por una matriz ortogonal preserva el producto interno y las normas

$$\begin{aligned}\langle \mathbf{v}, \mathbf{w} \rangle &= \langle A\mathbf{v}, A\mathbf{w} \rangle \\ \|A\mathbf{x}\| &= \|\mathbf{x}\|.\end{aligned}$$

Como consecuencia de la proposición anterior, el mapeo de vectores multiplicado por una matriz ortogonal preserva el ángulo entre los dos vectores. Además, la preservación de la norma implica que la multiplicación de un vector por una matriz ortogonal preserva la distancia entre el vector y el origen. Por lo tanto, podemos interpretar el mapeo $\mathbf{x} \rightarrow A\mathbf{x}$ para una matriz ortogonal A como la rotación o la reflexión en n dimensiones.

Definición 1.11 El producto de Kronecker de dos matrices $A \in \mathbb{R}^{m \times m}$ y $B \in \mathbb{R}^{n \times n}$ es

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1m}B \\ a_{21}B & a_{22}B & \cdots & a_{2m}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mm}B \end{pmatrix}$$

donde $a_{ij}B$ es la matriz que corresponde al producto de un escalar y una matriz.

El producto Kronecker es consistente con la definición anterior de producto externo de dos vectores $\mathbf{v} \otimes \mathbf{w}$

$$\mathbf{u} \otimes \mathbf{v} = \mathbf{uv}^T.$$

1.2 Rango y nulidad

Definición 1.12 El espacio lineal generado por los vectores $S = \{\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(n)}\}$ es el conjunto de todas las combinaciones lineales de vectores en S . Una base de un espacio lineal S es cualquier conjunto de vectores linealmente independientes que lo genere. La dimensión $\dim S$ de un espacio lineal S es el tamaño de su base.

Ejemplo 1.3 El espacio \mathbb{R}^n es generado por la base estándar $\mathbf{e}^{(i)}, i = 1, \dots, n$. Desde que los vectores de la base estándar son linealmente independientes, ellos forman una base para \mathbb{R}^n . Otras posible base para \mathbb{R}^n es $\{\mathbf{u}^{(i)} : i = 1, \dots, n\}$ donde $\mathbf{u}^{(i)}$ es definido como $u_j^{(i)} = 1$ si $j \leq i$ y 0 en otros casos.

Definición 1.13 El espacio columna de una matriz $A \in \mathbb{R}^{n \times m}$ es el espacio $\text{col}(A)$ generado por las columnas de A o de forma equivalente

$$\text{col}(A) = \{A\mathbf{v} : \mathbf{v} \in \mathbb{R}^{m \times 1}\} \subset \mathbb{R}^n.$$

Nos referimos a $\dim \text{col}(A)$ como el rango del espacio columna de A . El espacio fila de $A \in \mathbb{R}^{n \times m}$ es el espacio $\text{fila}(A)$ generado por las filas de A . El espacio nulo o kernel de $A \in \mathbb{R}^{n \times m}$ es

$$\text{nul}(A) = \{\mathbf{v} : A\mathbf{v} = \mathbf{0}\} \subset \mathbb{R}^m.$$

Proposición 1.6 Para una matriz A ,

$$\dim \text{col}(A) = \dim \text{fila}(A).$$

Se denota a este número como **rango**(A).

En efecto consideremos una matriz A con $\dim \text{col}(A) = r$ y colocamos los r vectores generando $\text{col}(A)$ como columnas de una matriz C . Dado que las columnas de A son una combinación lineal de las columnas de C tenemos $A = CR$ para alguna matriz R con r filas. Cada fila de $A = CR$ es una combinación lineal de las filas de R y por lo tanto el espacio fila de A es un subconjunto del espacio de fila de R cuya dimensión está acotada por r .

La desigualdad inversa se obtiene aplicando el mismo argumento pero a la matriz transpuesta A^T .

Definición 1.14 Sea una matriz $A \in \mathbb{R}^{n \times m}$, es una matriz de rango completo si $m \neq n$ y para el cuál $\text{rango}(A) = \min(n, m)$.

Proposición 1.7 $\text{rango}(AB) \leq \min(\text{rango } A, \text{rango } B)$

En efecto, desde que las columnas de AB son una combinación lineal de las columnas de A , $\text{rango}(AB) \leq \text{rango}(A)$. De manera similar, las filas de AB son combinaciones lineales de las filas de B , $\text{rango}(AB) \leq \text{rango}(B)$.

Proposición 1.8 Para matrices no singulares P, Q , tenemos

$$\text{rango}(PAQ) = \text{rango}(A).$$

Aplicando la proposición anterior, tenemos

$$\begin{aligned} \text{rango}(PAQ) &\leq \text{rango}(AQ) \leq \text{rango}(A), \\ \text{rango}(A) &= \text{rango}(IAI) = \text{rango}(P^{-1}PAQQ^{-1}) \leq \text{rango}(PAQ). \end{aligned}$$

Proposición 1.9 Para una matriz $A \in \mathbb{R}^{m \times n}$

$$\text{rango}(A) + \dim(\text{nul}(A)) = n.$$

Proposición 1.10

$$\text{rango}(A) = \text{rango}(A^T) = \text{rango}(AA^T) = \text{rango}(A^T A).$$

Desde que $Ax = 0$ implica que $A^T Ax = 0$ y desde que $A^T Ax$ implica $xA^T Ax = 0$, se tiene $Ax = 0$, de lo que se cumple $\text{nul}(A) = \text{nul}(A^T A)$. Desde que ambas matrices tienen el mismo número de columnas, las proposiciones previas implican que $\text{rango}(A) = \text{rango}(A^T A)$.

Aplicando el mismo argumento a A^T se muestra que $\text{rango}(A^T) = \text{rango}(AA^T)$. De la proposición 1.6 se tiene $\text{rango}(A) = \text{rango}(A^T)$. Lo que prueba el resultado.

1.3 Autovalores, autovectores, determinantes y traza

Definición 1.15 Sean $A \in \mathbb{R}^{n \times n}$ y $\mathbf{v} \in \mathbb{R}^n$ distinto de cero, se dice que \mathbf{v} es autovector de A si

$$A\mathbf{v} = \lambda\mathbf{v}, \quad \text{para algún } \lambda \in \mathbb{R}.$$

Se dice que λ es el autovalor de A , correspondiente a \mathbf{v} . El espectro de una matriz A es el conjunto de todos los autovalores.

A partir de la definición, podemos tener las siguientes observaciones

1. La definición implica que los autovalores y autovectores de A son soluciones de la ecuación vectorial $(A - \lambda I)\mathbf{v} = 0$.
2. El vector \mathbf{v} puede ser solución de $(A - \lambda I)\mathbf{v} = 0$ si $\dim \text{nul}(A - \lambda I) \geq 1$, implicando que $(A - \lambda I)$ es una matriz singular.
3. Si \mathbf{v} es un autovector de A , entonces lo es $c\mathbf{v}$ (con el mismo autovalor).

Definición 1.16 Consideremos el conjunto de los números $\{1, 2, \dots, n\}$. Una permutación es una determinada ordenación de sus elementos. Si $S = \{1, 2, \dots, n\}$ una permutación de orden n del conjunto S es una función biyectiva $\sigma : S \rightarrow S$, de manera que una permutación $\sigma : S \rightarrow S$ se puede escribir como

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$$

Por ejemplo, para el conjunto $S : \{1, 2, 3\}$, se tienen 6 permutaciones :

$$(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1).$$

El producto de dos permutaciones $\pi\sigma$ es definido por la función de composición $\pi \circ \sigma$. El conjunto de todas las permutaciones de orden n es denotado por Δ_n .

Definición 1.17 Un par $a, b \in \{1, \dots, n\}$ es una inversión de una permutación $\sigma \in \Delta_n$ si $a < b$ pero $\sigma(b) < \sigma(a)$.

En otras palabras, la función σ invierte el orden natural de a, b . El número de inversiones de una permutación se llama paridad. Las permutaciones con una paridad par se llaman pares y las permutaciones con una paridad impar se llaman impares.

El signo de una permutación σ es denotado por $s(\sigma)$, es igual a 1 si σ es par y -1 si σ es impar.

Definición 1.18 El determinante de una matriz cuadrada $A \in \mathbb{R}^{n \times n}$ es definida como

$$\det A = \sum_{\sigma \in \Delta_n} s(\sigma) A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{n,\sigma(n)}.$$

La definición anterior establece que el determinante es una suma de muchos términos, cada uno un producto de los elementos de la matriz de cada fila y con diferentes columnas. La suma alterna entre sumar y restar estos productos, dependiendo de la paridad de la permutación.

Ejemplo 1.4 El determinante de una matriz A de orden 2×2 es

$$\det A = a_{11}a_{22} - a_{12}a_{21}.$$

Hay dos precisamente dos permutaciones en Δ_2 ; la identidad $\sigma_1(\sigma_1(i) = i)$ y $\sigma_2(\sigma_2(1) = 2, \sigma_2(2) = 1)$. La permutación identidad tiene cero inversiones y por tanto es par. La permutación σ_2 tiene una inversión (el par $(1, 2)$ y por lo tanto es impar).

Ejemplo 1.5 El determinante de una A de orden 3×3 es

$$\begin{aligned} \det A = & a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{1,2}a_{2,3}a_{3,1} \\ & + a_{1,3}a_{2,1}a_{3,2} - a_{1,3}a_{2,2}a_{3,1}. \end{aligned}$$

El primer término corresponde a la permutación identidad, que es par. Los siguientes dos términos corresponde a permutaciones con una sola inversión ($(2, 3)$ en el primer caso y $(1, 2)$ en el segundo caso), haciéndolos impares. Los dos términos siguientes tienen dos inversiones que hacen pares ($(2, 3)$ y $(1, 3)$ en el primer caso y $(1, 2)$ y $(1, 3)$ en el segundo caso). El último término tiene tres inversiones que lo hacen impar ($(1, 2), (1, 3), (2, 3)$).

Proposición 1.11 El determinante de una matriz triangular es el producto de los elementos de la diagonal.

Proposición 1.12 Las siguientes propiedades se cumplen

1. $\det I = 1$.
2. $\det A$ es una función lineal del j -ésimo vector columna $\mathbf{v} = (A_{1j}, \dots, A_{nj})$, suponiendo que otras columnas se mantienen fijas.
3. Si A' es obtenida desde A intercambiando dos columnas, entonces $\det A = -\det A'$.
4. Si A tiene dos columnas iguales, entonces $\det A = 0$.
5. $\det(A) = \det(A^T)$.

Proposición 1.13 Para dos matrices cuadradas A, B tenemos

$$\begin{aligned} \det(BA) &= \det B \det A \\ \det(A^{-1}) &= (\det A)^{-1}. \end{aligned}$$

Proposición 1.14 Una matriz cuadrada es singular si y sólo si su determinante es cero.

Definición 1.19 El polinomio característico de una matriz cuadrada A es la función

$$f(\lambda) = \det(\lambda I - A), \quad \lambda \in \mathbb{R}.$$

1. Los autovalores son las raíces del polinomio característico

$$f(\lambda) = \det(\lambda I - A) = 0.$$

2. Si tenemos los autovalores $\lambda_1, \dots, \lambda_k$ son conocidos, obtenemos los autovectores, resolviendo el sistema de ecuaciones lineales

$$(A - \lambda_i I)\mathbf{v}^{(i)} = \mathbf{0}, \quad i = 1, \dots, k.$$

3. Desde que los autovalores $\lambda_1, \dots, \lambda_n$ son las raíces el polinomio característico $f(\lambda)$, puede ser escrito como el siguiente producto

$$f(\lambda) = \det(\lambda I - A) = \prod_i (\lambda - \lambda_i).$$

Definición 1.20 La traza de una matriz cuadrada $A \in \mathbb{R}^{n \times n}$ como

$$\text{Tr}(A) = \sum_{i=1}^n a_{ii}, \quad \text{donde } A = [a_{ij}]_{n \times n}.$$

Proposición 1.15 Las siguientes propiedades de la traza, se cumplen, $c \in \mathbb{R}$

$$\begin{aligned} \text{Tr}(A + B) &= \text{Tr}(A) + \text{Tr}(B) \\ \text{Tr}(AB) &= \text{Tr}(BA) \\ c \text{Tr}(A) &= \text{Tr}(cA) \end{aligned}$$

Definición 1.21 La norma de Frobenius de una matriz A es la norma Euclideana de un vector obtenido, concatenando las columnas de la matriz en un vector

$$\|A\|_F = \sqrt{\sum_i \sum_j a_{ij}^2}.$$

Desde que los elementos de la diagonal de AA^T , son la suma de cuadrados de las filas de A y los elementos de la diagonal de $A^T A$ son la suma de cuadrados de las columnas de A , tenemos

$$\|A\|_F = \sqrt{\text{Tr}(A^T A)} = \sqrt{\text{Tr}(AA^T)}.$$

Proposición 1.16 Para una matriz A y una matriz ortogonal U ,

$$\|UA\|_F = \|A\|_F.$$

Proposición 1.17 Si A es una matriz cuadrada con autovalores $\lambda_i, i = 1, \dots, n$

$$\begin{aligned}\text{Tr } A &= \sum_{i=1}^n \lambda_i \\ \det A &= \prod_{i=1}^n \lambda_i.\end{aligned}$$

En efecto, expresando el polinomio característico $f(\lambda)$ como producto $\prod_i (\lambda - \lambda_i)$, tenemos

$$0 = f(\lambda) = \prod_{i=1}^n (\lambda - \lambda_i) = \lambda^n - \lambda^{n-1} \sum_{i=1}^n \lambda_i + \dots + (-1)^n \prod_{i=1}^n \lambda_i.$$

Si multiplicamos una columna por -1 invertimos el signo del determinante. Se sigue entonces $\det A = (-1)^n \det(-A)$ y como $f(0) = \det(-A) = (-1)^n \prod_{i=1}^n \lambda_i$, se tiene que $\det A = \prod_{i=1}^n \lambda_i$.

Sustituyendo la definición del determinante en $\det(\lambda I - A)$, los términos de potencias λ^{n-1} resultan de una multiplicación de los términos diagonales $\prod_{i=1}^n (\lambda - a_{ii})$. Más específicamente, hay n términos conteniendo una potencia λ^{n-1} en la expansión del determinante $-a_{11}\lambda^{n-1}, \dots, -a_{nn}\lambda^{n-1}$.

Coleccionando, esos términos, conseguimos que el coeficiente asociado con λ^{n-1} en el polinomio característico es $-\text{Tr}(A) = -\sum_{i=1}^n a_{ii}$. Comparando esto al coeficiente de λ^{n-1} en la ecuación de arriba, tenemos $\text{Tr}(A) = \sum \lambda_i$.

Definición 1.22 Sean dos matrices A y B , se dicen que son similares, si existe una matriz invertible S tal que $S^{-1}AS = B$.

Proposición 1.18 Matrices similares tienen los mismos autovalores.

Definición 1.23 Si A es similar a una matriz diagonal, esto es, existe una matriz invertible S tal que $S^{-1}AS = D$ donde D es una matriz diagonal, entonces se dice que A es diagonalizable.

Proposición 1.19 (Teorema de Diagonalizabilidad) Una matriz $A \in \mathbb{R}^{m \times n}$ es diagonalizable si y sólo si tiene n autovectores independientes.

Supongamos que A es diagonalizable. Entonces $A = SDS^{-1}$, para una matriz diagonal D y una matriz invertible S . Supongamos que \mathbf{x}^0 es un vector. La ecuación $\mathbf{x}^{(t+1)} = A\mathbf{x}^{(t)}$ define los vectores $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}, \dots$. Entonces

$$\begin{aligned}\mathbf{x}^{(t)} &= \underbrace{AA \cdots A}_{t \text{ veces}} \mathbf{x}^{(0)} \\ &= (SDS^{-1})(SDS^{-1}) \cdots (SDS^{-1})\mathbf{x}^{(0)} \\ &= SD^t S^{-1} \mathbf{x}^{(0)}\end{aligned}$$

Sea $\mathbf{u}^{(t)}$ la representación de coordenadas de $\mathbf{x}^{(t)}$ en términos de las columnas de S . Entonces tenemos la ecuación $\mathbf{u}^{(t+1)} = D\mathbf{u}^{(t)}$. Por tanto

$$\begin{aligned}\mathbf{u}^{(t)} &= \underbrace{DD \cdots D}_{t \text{ veces}} \mathbf{u}^{(0)} \\ &= D^t \mathbf{u}^{(0)}\end{aligned}$$

$$\text{Si } D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \text{ entonces } D^T = \begin{pmatrix} \lambda_1^T & & \\ & \ddots & \\ & & \lambda_n^T \end{pmatrix}$$

Proposición 1.20 (Descomposición Espectral) Para una matriz A simétrica, tenemos

$$U^T A U = \mathbf{diag}(\lambda) \quad A = U \mathbf{diag}(\lambda) U^T,$$

donde λ es el vector de autovalores y U es una matriz ortogonal cuyas columnas son los autovectores correspondientes.

Proposición 1.21 Una matriz ortogonal cuadrada no es singular y tiene un determinante $+1$ ó -1 .

Proposición 1.22 Una matriz cuadrada es ortogonal si y sólo si $A^{-1} = A^T$. Además A es ortogonal si y sólo si A^T es ortogonal.

Toda transformación lineal T se realiza mediante una operación de multiplicación matricial: $T(\mathbf{x}) = A\mathbf{x}$. Si A es ortogonal, la aplicación $\mathbf{x} \rightarrow A\mathbf{x}$ puede interpretarse como una rotación o reflexión geométrica alrededor del eje y la aplicación $\mathbf{x} \rightarrow A^T \mathbf{x} = A^{-1} \mathbf{x}$ es la rotación o reflexión inversa.

Si A es diagonal, la aplicación $\mathbf{x} \rightarrow A\mathbf{x}$ puede interpretarse como estirar algunas dimensiones y comprimir otras dimensiones.

Aplicando la descomposición espectral a una matriz A simétrica, obtenemos una descomposición de A como un producto de tres matrices $U \mathbf{diag}(\lambda) U^T$. Esto implica que la transformación lineal $T(\mathbf{x}) = A\mathbf{x}$ puede ser vista como una secuencia de tres transformaciones lineales: la primera inicia una rotación o reflexión, la segunda escala las dimensiones y la tercera es otra rotación o reflexión.

Proposición 1.23 Para una matriz simétrica A , $\mathbf{rango}(A)$ es igual al número de valores propios distintos de cero.

Proposición 1.24 Para una matriz $A \in \mathbb{R}^{n \times n}$ simétrica tiene n autovectores ortonormales y $\mathbf{col}(A)$ es generado por los autovectores correspondientes a autovalores distintos de ceros.

1.4 Matrices definidas positivas

Definición 1.24 Una matriz cuadrada $A \in \mathbb{R}^{n \times n}$ es semidefinida positiva si

$$\mathbf{v}^T A \mathbf{v} \geq 0, \quad \forall \mathbf{v} \in \mathbb{R}^n$$

y definida positiva si la desigualdad se mantiene con igualdad sólo para vectores $\mathbf{v} = \mathbf{0}$. Una matriz cuadrada $A \in \mathbb{R}^{n \times n}$ es semidefinida negativa si

$$\mathbf{v}^T A \mathbf{v} \leq 0, \quad \forall \mathbf{v} \in \mathbb{R}^n$$

y definida negativa si la desigualdad se mantiene con igualdad sólo para vectores $\mathbf{v} = \mathbf{0}$

Proposición 1.25 Una matriz simétrica es semidefinida positiva si y sólo si todos los autovalores no son negativos. Es semidefinida positiva si y sólo si todos los autovalores no son positivos. Es definida positiva si y sólo si todos los valores propios son positivos. Es definida negativa si y sólo si todos los autovalores son negativos.

En efecto \mathbf{v} un vector arbitrario. Por la descomposición espectral, se tiene

$$\mathbf{v}^T A \mathbf{v} = (\mathbf{v}^T U) \mathbf{diag}(\lambda)(U^T \mathbf{v}) = \sum_{i=1}^n \lambda_i ([\mathbf{v}^T U]_i)^2$$

donde U es una matriz conteniendo los n autovectores ortogonales de A . La expresión anterior es no negativa para todo \mathbf{v} si y sólo si $\lambda_i \geq 0$ para todo $i = 1, \dots, n$. Los otros resultados se prueban de manera similar.

Proposición 1.26 Las matrices definitivas positivas y negativas positivas son necesariamente no singulares.

Proposición 1.27 Una matriz simétrica de rango r es semidefinida positiva si y sólo si existe una matriz cuadrada R de rango r tal que $A = R^T R$. Si A es definida positiva entonces R es singular.

Proposición 1.28 Si A es definida positiva, entonces A^{-1} lo es también.

Proposición 1.29 Los elementos de la diagonal de una matriz definida positiva, son todos positivos.

Proposición 1.30 Si A es definida positiva, entonces existe la matriz raíz cuadrada $A^{1/2}$ que cumple $A^{1/2} A^{1/2} = A$.

Sea A una matriz definida positiva con autovalores positivos. usando la descomposición espectral, tenemos

$$\begin{aligned} A &= U^T \mathbf{diag}(\lambda_1, \dots, \lambda_n) U \\ &= (U^T \mathbf{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}) U) (U^T \mathbf{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}) U) \\ &= A^{1/2} A^{1/2}. \end{aligned}$$

1.5 Algunas descomposiciones

Definición 1.25 Sea $v \in \mathbb{R}^n$ y P un subespacio de \mathbb{R}^n , se define la proyección ortogonal de v sobre P como $w = \text{proj}_P(v) \in P$ tal que

$$\|w - v\| \leq \|z - v\| \quad \text{para todo } z \in P.$$

Definición 1.26 Dado $\{u_1, u_2, \dots, u_m\}$ base de un subespacio $S \subset \mathbb{R}^n$ y sea $v \in \mathbb{R}^n$, entonces la proyección ortogonal de v sobre S es

$$\text{proj}_S(v) = U(U^T U)^{-1} U^T v \quad \text{donde } U = [u_1 \ \dots \ u_m]$$

$P = U(U^T U)^{-1} U^T$ se conoce como matriz de proyección.

Definición 1.27 (Descomposición ortogonal) Dado $\{u_1, u_2, \dots, u_m\}$ una base ortogonal de un subespacio $S \subset \mathbb{R}^n$ y sea $v \in \mathbb{R}^n$, entonces la proyección ortogonal de v sobre S es

$$\text{proj}_S(v) = \frac{v \cdot u_1}{\|u_1\|^2} u_1 + \frac{v \cdot u_2}{\|u_2\|^2} u_2 + \dots + \frac{v \cdot u_m}{\|u_m\|^2} u_m.$$

Si $\{u_1, u_2, \dots, u_m\}$ es una base ortonormal entonces:

$$\text{proj}_S(v) = (v \cdot u_1)u_1 + (v \cdot u_2)u_2 + \dots + (v \cdot u_m)u_m$$

Por propiedad $\text{proj}_S(v) = U(U^T U)^{-1} U^T v$, donde $U = [u_1 \ \dots \ u_m]$.

Dado que el conjunto $\{u_1, u_2, \dots, u_m\}$ es ortogonal entonces $u_i^T u_j = 0$ para todo $i \neq j$, luego

$$(U^T U)^{-1} = \left(\begin{bmatrix} u_1^T \\ \vdots \\ u_m^T \end{bmatrix} [u_1 \ \dots \ u_m] \right)^{-1} = \left(\begin{bmatrix} \|u_1\|^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \|u_m\|^2 \end{bmatrix} \right)^{-1} = \begin{bmatrix} \frac{1}{\|u_1\|^2} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\|u_m\|^2} \end{bmatrix}$$

Luego,

$$\text{proj}_S(v) = [u_1 \ \dots \ u_m] \begin{bmatrix} \frac{1}{\|u_1\|^2} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\|u_m\|^2} \end{bmatrix} \begin{bmatrix} u_1^T \\ \vdots \\ u_m^T \end{bmatrix} v = [u_1 \ \dots \ u_m] \begin{bmatrix} \frac{u_1^T v}{\|u_1\|^2} \\ \vdots \\ \frac{u_m^T v}{\|u_m\|^2} \end{bmatrix}$$

Por tanto $\text{proj}_S(v) = \sum_{i=1}^m \frac{v \cdot u_i}{\|u_i\|^2} u_i$. Si $\{u_1, \dots, u_m\}$ es un conjunto ortogonal entonces $\|u_i\| = 1$, por lo

que $\text{proj}_S(v) = \sum_{i=1}^m (v \cdot u_i) u_i$.

Proposición 1.31 Si \mathbf{v} es un autovector, con valor propio $\lambda \neq 0$, de $A^T A$, luego $A\mathbf{v}$ es un autovector con el mismo valor propio λ de AA^T .

En efecto sea \mathbf{v} es un autovector de $A^T A$ y $\lambda \neq 0$ su autovalor correspondiente, entonces $A^T A\mathbf{v} = \lambda\mathbf{v}$. Luego multiplicando por A a la izquierda, a la anterior expresión

$$\begin{aligned} A(A^T A\mathbf{v}) &= A(\lambda\mathbf{v}) \\ \Rightarrow AA^T(A\mathbf{v}) &= \lambda(A\mathbf{v}). \end{aligned}$$

Por lo que $A\mathbf{v}$ es autovector de AA^T y $\lambda \neq 0$ su autovalor correspondiente.

Proposición 1.32 Sea $A \in \mathbb{R}^{n \times n}$, entonces $A^T A$ no tiene autovalores negativos.

Para verificar este resultado, sea \mathbf{u} un autovector de $A^T A$ y λ su autovalor correspondiente, entonces $A^T A\mathbf{u} = \lambda\mathbf{u}$, luego

$$\begin{aligned} \mathbf{u}^T(A^T A\mathbf{u}) &= \mathbf{u}^T(\lambda\mathbf{u}) \\ (\mathbf{u}^T A^T)A\mathbf{u} &= \lambda\mathbf{u}^T\mathbf{u} \\ (A\mathbf{u})^T A\mathbf{u} &= \lambda\|\mathbf{u}\|^2 \\ \|A\mathbf{u}\|^2 &= \lambda\|\mathbf{u}\|^2 \\ 0 \leq \frac{\|A\mathbf{u}\|^2}{\|\mathbf{u}\|^2} &= \lambda \end{aligned}$$

Definición 1.28 Si tenemos una matriz $A \in \mathbb{R}^{n \times m}$, se definen los valores singulares de A como las raíces cuadradas de los eigenvalores de $A^T A$ y de AA^T .

Proposición 1.33 Para una matriz $A \in \mathbb{R}^{n \times m}$, tenemos

$$A = U\Sigma V^T, \quad \text{donde}$$

- $U \in \mathbb{R}^{n \times n}$ es una matriz ortogonal cuyas columnas son los autovectores de AA^T .
- $V \in \mathbb{R}^{m \times m}$ es una matriz ortogonal cuyas columnas son los autovectores de $A^T A$.
- $\Sigma \in \mathbb{R}^{n \times m}$ es una matriz de ceros, excepto, los primeros r elementos de la diagonal σ_i , $i = 1, \dots, r$, es decir los valores singulares.

Se asume en este resultado que los valores singulares se ordenan en forma descendente y los autovectores se ordenan según el orden descendente de sus autovalores.

$$A = U\Sigma V^T, \quad \text{donde} \quad \Sigma = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix}, \quad \sigma_1 \geq \dots \geq \sigma_m \geq 0,$$

Cuando A tiene rango completo, los valores singulares son estrictamente positivos y

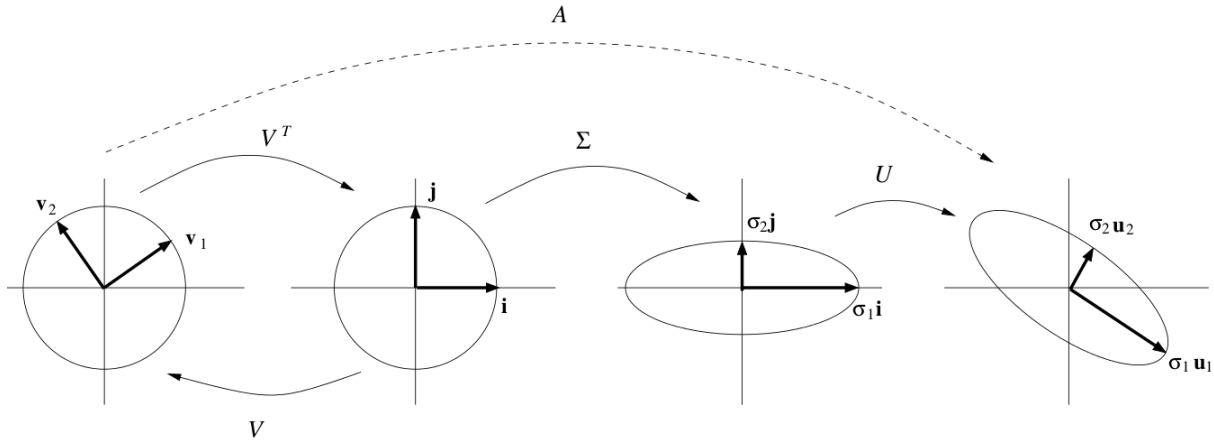
$$A = U \begin{pmatrix} \Sigma & 0 \end{pmatrix} V^T \quad \text{o} \quad A = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^T$$

Corolario 1.1 Si denotamos el valor singular por σ , se cumple

$$\|A\|_F = \|\sigma\|.$$

La descripción intuitiva de la descomposición espectral indica que cualquier transformación lineal $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ correspondiente a una multiplicación por una matriz simétrica es equivalente a una rotación o una reflexión, seguida de un escalamiento de los ejes y seguida de una rotación o reflexión inversa.

La técnica SVD generaliza esta interpretación a matrices no simétricas e incluso no cuadradas. Cualquier transformación lineal $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$ se puede descomponer como una secuencia de una rotación o reflexión, seguida por un escalamiento de los ejes, seguida por otra rotación o escalamiento.



Geoméricamente, esto implica que la aplicación $x \rightarrow Ax$ transforma una esfera en una elipse. Las orientaciones de los ejes de la elipse son proporcionadas por las columnas de U , V y el alargamiento de los diferentes ejes de la elipse se determinan con los valores singulares, como indica la figura anterior.

2 Referencias

1. Coding the Matrix, Philip Klein, Newtonian Press, 2013.
2. Probability, The Analysis of Data, volumen 1 Guy Lebanon.
3. Linear Algebra Done Right Sheldon Axler 2015.