Repaso de Álgebra lineal 2

Reconocimiento de patrones

Gamaliel Moreno

gamalielmch@uaz.edu.mx
http://pds.uaz.edu.mx/

Enero-Julio 2021

Contenido

Inversa de una matriz

- Una matriz cuadrada $\boldsymbol{A}$ es invertible (o no singular) si existe otra matriz $\boldsymbol{A}^{-1}$ tal que

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$$

- Caso contrario, la matriz se dice singular (o no invertible)
- Una matriz es ortogonal si se cumple $\boldsymbol{A}^{-1} = \boldsymbol{A}^{T}$, es decir, si sus vectores (fila y columna) son ortonormales entre sí.
- Una matriz ortogonal no cambia la norma l_2 de un vector $||\boldsymbol{A}\boldsymbol{x}||_2 = ||\boldsymbol{x}||_2$

Determinantes de una matriz 1

- La determinante de una matriz cuadrada es un valor escalar denotado como $|\boldsymbol{A}|$ o $det\boldsymbol{A}$
- Interpretación: Sea $\boldsymbol{\alpha}$ un vector con $0 \leq \alpha_i \leq 1$. Todos los vectores que cumplen esta condición conforman un hipercubo en el espacio n dimensional. La determinante indica el volumen de ese hipercubo transformado por la matriz $\boldsymbol{A}$
- La determinant de una matriz $2x2$ se calcula con

$$|\boldsymbol{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Determinantes de una matriz 1

- Para una matriz de mayor orden la determinante es

$$|\boldsymbol{A}| = \sum_{j=1}^{n} a_{fj} |\boldsymbol{A}_{fj}|$$

donde $\boldsymbol{A}_{fj}$ es la matriz adjunta de a_{fj} dada por $(-1)^{f+j}\boldsymbol{M}_{fj}$ y $\boldsymbol{M}_{fj}$ es el menor complementario de a_{fj}, es decir, una matriz $(n-1)\times(n-1)$ obtenida eliminando la fila f y la columna j de la matriz $\boldsymbol{A}$

Propiedades de la determinante

Sea $\boldsymbol{A}$ una matriz cuadrada de $n \times n$

- $|s\boldsymbol{A}| = s^n|\boldsymbol{A}|$
- $|\boldsymbol{I}| = 1$
- Distributividad: $|\boldsymbol{AB}| = |\boldsymbol{A}||\boldsymbol{B}|$
- $|\boldsymbol{I}| = 1 = |\boldsymbol{AA}^{-1}| = |\boldsymbol{A}||\boldsymbol{A}^{-1}| \Rightarrow |\boldsymbol{A}^{-1}| = 1/|\boldsymbol{A}|$
- $|\boldsymbol{A}| = |\boldsymbol{A}^T|$
- $|\boldsymbol{A}| = 0$ sí y sólo sí $\boldsymbol{A}$ es singular

Normas

- Informalmente, la norma de $\|\boldsymbol{x}\|$ es una medida de la longitud del vector $\boldsymbol{x} \in \mathbb{R}^n$
- Formalmente, la norma es un operador $\| \cdot \| : \mathbb{R}^n \to \mathbb{R}$ que satisface 4 propiedades:
  1. $\forall \boldsymbol{x} \in \mathbb{R}^n, \|\boldsymbol{x}\| \geq 0$ (no negatividad)
  2. $\|\boldsymbol{x}\| = 0 \Leftrightarrow \boldsymbol{x} = 0$ (definitud)
  3. $\forall \boldsymbol{x} \in \mathbb{R}^n, a \in \mathbb{R}, \|a\boldsymbol{x}\| = |a|\|\boldsymbol{x}\|$ (Homogeneidad)
  4. $\forall \boldsymbol{x} \in \mathbb{R}^n, \|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ (desigualdad de Minkowski)

Ejemplo de normas

- Normas ℓ_p de Minkowski o normas p para vectores $\boldsymbol{x}\ in \mathbb{R}^n$:

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

  casos particulares
- Formalmente, la norma es un operador $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ que satisface 4 propiedades:
  - Norma ℓ_2 o euclidiana: $\|\boldsymbol{x}\|_2 = \sqrt{\boldsymbol{x}^T\boldsymbol{x}} = \sqrt{\sum_{i=1}^{n} x_i^2}$
  - Norma ℓ_1 o de bloques de ciudad:$\|\boldsymbol{x}\|_1 = \sum_{i=2}^{n} |x_i|$
  - Norma ℓ_∞ : $\|\boldsymbol{x}\|_\infty = max|x_i|$
- Norma de Frobenius para matrices $\boldsymbol{A} \in \mathbb{R}^{m\times n}$

$$\|\boldsymbol{A}\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2} = \sqrt{tr(\boldsymbol{A}^T\boldsymbol{A})}$$

Independencia lineal

- El conjunto $\{\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_m}\} \subset \mathbb{R}$ es linealmente independiente si ningún vector puede expresarse como combinación lineal de los otros vectores. Este es el caso sólo si

$$\sum_{i=1}^{m} \alpha_i \boldsymbol{x_i} = 0 \Leftrightarrow \alpha_1 = \alpha_2 = \ldots = \alpha_m = 0$$

- Caso contrario se dice que los vectores son linealmente dependientes.

$$\boldsymbol{x}_j = \sum_{i=1,\ldots,m, i \neq j} \alpha_i \boldsymbol{x}_i$$

Bases

- Si $\mathcal{B}$ es un conjunto generador de un espacio lineal $\mathbb{V}$, $\mathcal{B}$ es una base de $\mathbb{V}$ si y sólo si todos sus vectores son linealmente independientes.
- El número de vectores en la base $n = |\mathcal{B}|$ es igual a la dimensión de $\mathbb{V}$
- Un conjunto generador de $\mathbb{V}$ requiere al menos n vectores
- Un conjunto generador linealmente independiente puede tener a lo sumo n vectores

Sistemas de ecuaciones en notación matricial 1

Todo sistema de ecuaciones lineales se puede expresa de la forma

$$\boldsymbol{Ax} = \boldsymbol{b}$$

que representa a

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

o en forma tradicional

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \quad + \quad \vdots \quad + \; \vdots \; + \quad \vdots \quad &= \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

Sistemas de ecuaciones en notación matricial 2

Si $\boldsymbol{b}$ está en el alcance columna de $\boldsymbol{A}$ entonces el sistema tiene solución:

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

El rango columna de $\boldsymbol{A}$ (rank) es el máximo número de columnas linealmente independientes, y es igual a la dimensión del alcance columna (range) de $\boldsymbol{A}$

El espacio nulo de $\boldsymbol{A}$ es el conjunto de vectores $\boldsymbol{z}$ tales que $\boldsymbol{A}\boldsymbol{z} = 0$. Su dimensión se llama nulidad.

La nulidad más el rango es igual al número de filas $\boldsymbol{A}$-

De forma dual, para el sistema $\boldsymbol{x}^T\boldsymbol{A}$ se define el espacio fila como la combinación lineal de las filas de $\boldsymbol{A}$, que tiene su alcance y rango fila.

Para $\boldsymbol{A} \in \mathbb{R}^{m \times c}$, se cumple que el rango columna y el rango fila son siempre iguales, y se denota como $rg\boldsymbol{A}$.

Propiedades del rango

- Para $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $rg(\boldsymbol{A}) \leq min(m,n)$. Si $rg(\boldsymbol{A}) = min(m,n)$, entonces se dice que $\boldsymbol{A}$ tiene rango completo.
- Solo las matrices de rango completo son invertibles
- $rg(\boldsymbol{A}) = rg(\boldsymbol{A}^T)$
- $rg(\boldsymbol{AB}) \leq min(rg(\boldsymbol{A}), rg(\boldsymbol{B}))$
- $rg(\boldsymbol{A} + \boldsymbol{B}) \leq rg(\boldsymbol{A}) + rg(\boldsymbol{B})$

Forma cuadrática y matrices (semi) definida positivas

- La forma cuadrática es la expresión escalar $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$, calculada con:

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \sum_{i=1}^{n} x_i(\boldsymbol{A}\boldsymbol{x})_i = \sum_{i=1}^{n} x_i \left(\sum_{j=1}^{n} a_{ij}x_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j$$

- como la forma cuadrática es escalar y $s = s^T$ entonces

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = (\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x})^T = \boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{x}$$

- como los tres términos anteriores son iguales, entonces

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \frac{1}{2}(\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{x}) = \boldsymbol{x}^T\left(\frac{1}{2}(\boldsymbol{A} + \boldsymbol{A}^T)\right)\boldsymbol{x}$$

es decir, el valor solo depende de la componente simétrica de $\boldsymbol{A}$.

Positividad o negatividad (semi)definida

La matriz simétrica $\boldsymbol{A} \in \mathbb{S}^n$ es

- Positiva definida si $\forall \boldsymbol{x} \in \mathbb{R}^n \setminus 0 \quad \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} > 0$.
  Notación: $\boldsymbol{A} \succ 0 (o \boldsymbol{A} > 0)$.
  Todas las matrices positivas definidas: $\mathbb{S}^n_{++}$
- Positiva semidefinida $\forall \boldsymbol{x} \in \mathbb{R}^n \quad \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} \geq 0$
  Notación: $\boldsymbol{A} \succeq 0 (o \boldsymbol{A} \geq 0)$.
  Todas las matrices positivas definidas: $\mathbb{S}^n_{+}$
- Negativa definida si $\forall \boldsymbol{x} \in \mathbb{R}^n \setminus 0 \quad \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} < 0$.
  Notación: $\boldsymbol{A} \prec 0 (o \boldsymbol{A} < 0)$.
  Todas las matrices positivas definidas: $\mathbb{S}^n_{--}$
- Negativa semidefinida si $\forall \boldsymbol{x} \in \mathbb{R}^n \quad \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} \geq 0$.
  Notación: $\boldsymbol{A} \preceq 0 (o \boldsymbol{A} \leq 0)$.
  Todas las matrices positivas definidas: $\mathbb{S}^n_{-}$
- Indefinida en cualquier otro caso

Características de matrices (semi)definida

- Si $\boldsymbol{A} \in \mathbb{S}^n_{++}$ entonces, $-\boldsymbol{A} \in \mathbb{S}^n_{--}$
- Si $\boldsymbol{A} \in \mathbb{S}^n_{+}$ entonces, $-\boldsymbol{A} \in \mathbb{S}^n_{-}$
- Todas las matrices positivas (o negativas) definidas sin de rango completo.
- Sea $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ una matriz cualquiera
  - $\boldsymbol{A}^T\boldsymbol{A}$ se denomina matriz Gram
  - $\boldsymbol{A}^T\boldsymbol{A}$ es siempre positiva semidefinida
  - Si $m \geq n$ y $\boldsymbol{A}$ es de rango completo entonces $\boldsymbol{A}^T\boldsymbol{A}$ es positiva definida

Eigenvalores y eigenvectores

- También llamados vectores propios, autovectores o vectores característicos
- Dada una matriz cuadrada $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, se dice que $\lambda \in \mathbb{C}$ es un eigenvalor de $\boldsymbol{A}$ y $\boldsymbol{x} \in \mathbb{C}^n$ es su correspondiente eigenvector si

$$\boldsymbol{Ax} = \lambda \boldsymbol{x}, \quad \boldsymbol{x} \neq 0$$

- Interpretación: transformación de eigenvector $\boldsymbol{x}$ con $\boldsymbol{A}$ solo cambia su magnitud con un factor λ
- Cualquier escalamiento de un eigenvector es también eigenvector

$$\boldsymbol{A}(c\boldsymbol{x}) = c(\boldsymbol{Ax}) = c\lambda\boldsymbol{x} = \lambda(c\boldsymbol{x})$$

se usan eigenvectores normalizados ($\|\boldsymbol{x}\|_2 = 1$)

Eigenvalores

- Multiplicando por matriz identidad no modifica nada

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} &= \lambda \boldsymbol{I}\boldsymbol{x} \\ (\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x} &= 0 \end{aligned}$$

  que tiene como solución trivial $\boldsymbol{x} = 0$
- Otras soluciones con $\boldsymbol{x} \neq 0$ posibles solo si $\boldsymbol{A} - \lambda \boldsymbol{I}$ es singular (tiene nulidad mayor que cero), por lo que

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

Eigenvalores

Puesto que

$$\boldsymbol{A}-\lambda\boldsymbol{I}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}-\begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \lambda \\ 0 & 0 & \cdots & 0 \end{bmatrix}=$$

$$\begin{bmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{bmatrix}$$

Eigenvalores

- La ecuación $det|\boldsymbol{A} - \lambda \boldsymbol{I} = 0$ es entonces

$$det \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0$$

  que produce un polinomio de orden n en términos de λ.
- Existen por tanto n soluciones (posiblemente complejas)
- Problema: ecuación mal condicionado (se usan otros métodos de cálculode λ)

Eigenvalores

- Para encontrar uno de los eigenvalores λ_i se plantea un sistema de ecuaciones para encontrar el eigenvector correspondiente

$$(\boldsymbol{A} - \lambda_i \boldsymbol{I})\boldsymbol{x}_i = 0$$

- Nótese que soluciones λ pueden ser complejas o tener multiplicidad mayor a 1. Esto implica que puede
  - No existan eigenvectores
  - Existen menos eigenvectores que la dimensión del espacio
  - Nunca van a existir más eigenvectores que la dimensión del espacio

Propiedades de eigenvalores y eigenvectores

- $tr\boldsymbol{A} = \sum_i^n \lambda_i$
- $|\boldsymbol{A}| \prod_i^n \lambda_i$
- $rg\boldsymbol{A} = |\{\lambda|\lambda \neq 0\}|$
- Si $\exists \boldsymbol{A}^{-1} \implies \boldsymbol{A}^{-1}\boldsymbol{x}_i = (1/\lambda_i)\boldsymbol{x}_i$
- Los eigenvalores de $\boldsymbol{D} = diag(d_1, \ldots, d_n)$ son $\lambda_i = d_i$

Expresión simultánea

- Todos los eigenvectores se expresan simultáneament con

$$\boldsymbol{AX} = \boldsymbol{X\Lambda}$$

con

$$\boldsymbol{X} \in \mathbb{R}^{n\times n} = \begin{bmatrix} \boxed{\boldsymbol{x_1}} & \cdots & \boxed{\boldsymbol{x_n}} \end{bmatrix}$$

$$\boldsymbol{\Lambda} = diag(\lambda_1, \ldots, \lambda n) = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

Diagonalización

- Si los eigenvectores de $\boldsymbol{A}$ son linealmente independientes, entonces $\boldsymbol{X}$ es invertible. En ese caso

$$\begin{aligned} \boldsymbol{AX} &= \boldsymbol{X\Lambda} \\ \boldsymbol{AXX}^{-1} &= \boldsymbol{X\Lambda X}^{-1} \\ \boldsymbol{A} &= \boldsymbol{X\Lambda X}^{-1} \end{aligned}$$

- En ese caso, la metriz se denomina diagonalizable

$$\begin{aligned} \boldsymbol{AX} &= \boldsymbol{X\Lambda} \\ \boldsymbol{X}^{-1}\boldsymbol{AX} &= \boldsymbol{X}^{-1}\boldsymbol{X\Lambda} \\ \boldsymbol{X}^{-1}\boldsymbol{AX} &= \boldsymbol{\Lambda} \end{aligned}$$