Análisis numérico

Clase 3: Preliminares, algoritmos y computación

Joaquin Cavieres

Instituto de Estadística, Universidad de Valparaíso

Outline

Transpuesta y traza de sumas y productos

Si $\boldsymbol{A}+\boldsymbol{B}$ estan definidas (por ejemplo, tienen el mismo orden), entonces $(\boldsymbol{A}+\boldsymbol{B})^T=\boldsymbol{A}^T+\boldsymbol{B}^T$ y $\mathrm{tr}(\boldsymbol{A}+\boldsymbol{B})=tr(\boldsymbol{A})+tr(\boldsymbol{B})$. Si $\boldsymbol{A}$ es de dimensión $m\times n$ y $\boldsymbol{B}$ es de orden $n\times p$, entonces $\boldsymbol{AB}$ es $m\times n\times n\times p\equiv m\times p$, así que $(\boldsymbol{AB})^T$ es de orden $p\times m$. Por lo anterior entonces es fácil demostrar que $(\boldsymbol{AB})^T=\boldsymbol{B}^T\boldsymbol{A}^T$:

$$(\boldsymbol{AB})^T=((a_{ij})(b_{ij}))^T=(\sum_{j=1}^{n}a_{ij}b_{jk})^T=(\sum_{j=1}^{n}a_{kj}b_{bji})=(b_{jk})^T(a_{ij})^T=\boldsymbol{B}^T\boldsymbol{A}^T$$

Note que $\boldsymbol{A}^T$ es $n\times m$ y $\boldsymbol{B}^T$ es $p\times n$ así que el producto $\boldsymbol{A}^T\boldsymbol{B}^T$ no está definido pero si está definido el producto $\boldsymbol{B}^T\boldsymbol{A}^T$.

La $\mathrm{tr}(\boldsymbol{A})=\mathrm{tr}(\boldsymbol{A}^T)$ y si $\boldsymbol{A}$ y $\boldsymbol{B}$ son matrices de orden $m\times n$ y $n\times m$ entonces $\mathrm{tr}(\boldsymbol{AB})=\mathrm{tr}\ (\boldsymbol{BA})$ por que la $\mathrm{tr}(\boldsymbol{AB})=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}b_{ij}=\sum_{j=1}^{n}\sum_{i=1}^{m}a_{ij}b_{ij}=\mathrm{tr}(\boldsymbol{BA})$.

Matrices especiales

Matrices ortogonales

Una matriz cuadrada $\boldsymbol{A}$ de orden $p \times p$ es ortogonal si $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{I}_p$. Note que para una matriz cuadrada si $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}_p$ y si $\boldsymbol{A}$ es no singular, entonces necesariamente nosotros tenemos $\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{I}_p$ ya que $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}_p$ entonces $(\boldsymbol{A}^T)^{-1}\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{A}^T = (\boldsymbol{A}^T)^{-1}\boldsymbol{I}_p\boldsymbol{A}^T = \boldsymbol{I}_p$. También $\boldsymbol{A}^{-1} = \boldsymbol{A}^T$ y si $\boldsymbol{A}$ es ortogonal entonces $\boldsymbol{A}^T$ también es ortogonal.

Si $\boldsymbol{A}$ y $\boldsymbol{B}$ son ortogonales y ambas matrices con dimensión $p \times p$ entonces $\boldsymbol{A}\boldsymbol{B}$ es ortogonal, ya que $(\boldsymbol{A}\boldsymbol{B})^T\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}^T\boldsymbol{I}_p\boldsymbol{B} = \boldsymbol{B}^T\boldsymbol{B} = \boldsymbol{I}_p$.

Es posible tener una matrix $\boldsymbol{B}$ de orden $m \times n$ tal que $\boldsymbol{B}^T\boldsymbol{B} = \boldsymbol{I}_n$ pero $\boldsymbol{B}\boldsymbol{B}^T \neq \boldsymbol{I}_m$.

Matrices ortogonales

Observación

Dos matrices confortables $\boldsymbol{U}$ y $\boldsymbol{V}$ son llamadas a veces ortogonales si $\boldsymbol{UV} = 0$. Aquí se esta usando el termino "ortogonal" en el mismo sentido que si dos vectores son ortogonales. Es más adecuado decir que $\boldsymbol{U}$ es ortogonal a $\boldsymbol{V}$. De otra manera, si se dice que $\boldsymbol{U}$ y $\boldsymbol{V}$ son ortogonales, esto podría significar que ambas son matrices ortogonales.

Matrices

Matrices normales

Una matriz $p \times p$ es normal si $\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{A}^T\boldsymbol{A}$. Claramente todas las matrices simétricas y ortogonales son matrices normales.

Matrices idempotentes

Una matriz $\boldsymbol{A}$ de orden $p \times p$ es idempotente sii $\boldsymbol{A}^2 = \boldsymbol{A}$. Claramente $\boldsymbol{I}_p$ y $0_{p\times p}$ son idempotentes y también $\boldsymbol{x}\boldsymbol{x}^T$ con $\boldsymbol{x}\boldsymbol{x}^T = 1$.

Las matrices idempotentes juegan un rol clave en las estadísticas porque pueden considerarse proyecciones y, de hecho, también se denominan matrices de proyección.

Matrices

Matrices unipotentes

Una matriz $\boldsymbol{A}$ de orden $n \times n$ tal que $\boldsymbol{A}^2 = \boldsymbol{I}_n$ es uninpotente. Ejemplos de este tipo de matrices puede ser todas las matrices identidad y por ejemplo $\boldsymbol{A} = \begin{bmatrix} 1 & x \\ 0 & -1 \end{bmatrix}$

Matrices similares

Dos matrices son $\boldsymbol{A}$ y $\boldsymbol{B}$ se dicen similares si hay una matriz no singular $\boldsymbol{C}$ tal que $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C} = \boldsymbol{B}$.

Introducción

Una matriz $\boldsymbol{X}$ de orden $m \times n$ tiene n columnas, $\boldsymbol{x}_1, ..., \boldsymbol{x}_n$ (llamados "vectores columnas" de largo m) y m filas (llamadas "vectores filas" de largo n). Se dice que $\boldsymbol{x}_1$ y $\boldsymbol{x}_2$ son linealmente independiente si $a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 = 0$, donde a_1 y a_2 son números reales. Esto implica que $a_1 = a_2 = 0$. Un conjunto de $\boldsymbol{x}_1, ..., \boldsymbol{x}_n$ es linealmente independiente si $\sum a_i\boldsymbol{x}_i = 0$, lo que implica que todos los a_i00, o en otras palabras, ellos son linealmente independientes si no hay combinaciones lineales no triviales de ellos que son iguales a 0.

Definición

El rango columna de $\boldsymbol{X}$ es el máximo número de columnas linealmente independientes de $\boldsymbol{X}$. El rango fila de $\boldsymbol{X}$ es el máximo número de filas linealmente independientes de $\boldsymbol{X}$. El rango fila de $\boldsymbol{X}$ es el mismo que el rango columna de $\boldsymbol{X}^T$.

Existe un teorema que no es sencillo de probar el cual dice que el rango fila y el rango columna son iguales, así nosotros podemos hablar sobre el **rango** de $\boldsymbol{X}$ (expresado como $\rho(\boldsymbol{X})$) sin especificar si corresponde a la fila o a la columna. Por lo anterior entonces $\rho(\boldsymbol{A}) = \rho(\boldsymbol{A}^T)$ y además $\rho(\boldsymbol{A}\boldsymbol{A}^T) = \rho(\boldsymbol{A}^T\boldsymbol{A}) = \rho(\boldsymbol{A}) = \rho(\boldsymbol{A}^T)$ [2].

Rango de matrices

De acuerdo a la definición anterior, el rango columna de $\boldsymbol{X} \leq n$ y el rango fila de $\boldsymbol{X} \leq m$, así podemos asumir que $\rho(\boldsymbol{X}) \leq \min(m, n)$. El rango de una matriz 0 es cero ($\rho(\boldsymbol{X}) = 0$) solamente si $\boldsymbol{X} = 0$.

Rango de matrices

Recordatorio:

Rango

Número de filas o columnas que son linealmente independientes.

Rango de matrices

Recordatorio:

Rango

Número de filas o columnas que son linealmente independientes.

¿Que significa que sean linealmente independientes?

Ejemplo 1:

Se tiene a $\boldsymbol{A} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, un vector en $\mathbb{R}^2$ de orden $m \times 1$, por tanto si multiplicamos al escalar 0 con el vector $\boldsymbol{A}$ obtenemos el vector $\boldsymbol{B} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, el cual es una combinación trivial.

Una combinación trivial es aquella en que todos los escalares que multiplican a un vector (o vectores) son iguales a 0.

Rango de matrices

Para que no exista combinación trivial basta que al menos 1 escalar sea distinto de 0 y aún así el resultado de la sumatoria sea igual a 0. Esta combinación es llamada "combinación no trivial"

Definición: dependencia lineal

Si tenemos n vectores, entonces estos son linealmente dependientes si existe al menos una combinación no trivial en que el resultado de la sumatoria escalar*vector sea 0, esto es $\alpha_1 * v1 + \alpha_2 * v2+, ..., \alpha_k * vk = 0$.

Se tiene al vector $\boldsymbol{U} = \begin{bmatrix} 3 \\ 4 \\ -2 \end{bmatrix}$ y al vector $\boldsymbol{V} = \begin{bmatrix} -9 \\ -12 \\ 6 \end{bmatrix}$. ¿Son linealmente dependientes?

Ya sabemos que la forma trivial sería multiplicar cada vector por 0 para así obtener el vector de $\vec{0}$ en $\mathbb{R}^3$. Pero, ¿existe una forma no trivial?

Definición: dependencia lineal

Si tenemos n vectores, entonces estos son linealmente dependientes si existe al menos una combinación no trivial en que el resultado de la sumatoria escalar*vector sea 0, esto es $\alpha_1 * v1 + \alpha_2 * v2 +, ..., \alpha_k * vk = 0$.

Se tiene al vector $\boldsymbol{U} = \begin{bmatrix} 3 \\ 4 \\ -2 \end{bmatrix}$ y al vector $\boldsymbol{V} = \begin{bmatrix} -9 \\ -12 \\ 6 \end{bmatrix}$. ¿Son linealmente dependientes?

Una de las formas sería multiplicar el escalar 3 y 1 con los respectivos vectores $\boldsymbol{U}$ y $\boldsymbol{V}$.

Definición: independencia lineal

Si tenemos n vectores, entonces estos son linealmente independientes si existe una única solución trivial en que el resultado de la sumatoria escalar*vector sea 0, esto es $\alpha_1 * v1 + \alpha_2 * v2+, ..., \alpha_k * vk = 0$.

Se tiene al vector $\boldsymbol{U} = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$, al vector $\boldsymbol{V} = \begin{bmatrix} -3 \\ -2 \\ 8 \end{bmatrix}$ y al vector $\boldsymbol{Z} = \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix}$.

¿Son linealmente dependientes o indepentientes?

Rango de matrices

Se tiene al vector $\boldsymbol{U} = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$, al vector $\boldsymbol{V} = \begin{bmatrix} -3 \\ -2 \\ 8 \end{bmatrix}$ y al vector $\boldsymbol{Z} = \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix}$.
¿Son linealmente dependientes o indepentientes?

Utilización de metódos como el de Gauss-Jordan o el de Gauss u otro método iterativo.

Rango de matrices

Resumen

- Si hay infinitas soluciones (sistema homogéneo), y tenemos una solución no trivial entonces, los vectores son linealmente depentientes.
- Si hay una solución única trivial entonces los vectores son linealmente independientes.

Ejemplo 2:

Se tiene a $\boldsymbol{X} = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$, una matriz de orden 2×3, por tanto el rango $\rho(\boldsymbol{X}) \leq \min(2,3) = 2$, así $\rho(\boldsymbol{X})$ puede ser 1 o 2. Si $\rho(\boldsymbol{X}) = 1$ entonces las filas de $\boldsymbol{X}$ son linealmente independientes, por ejemplo: existen constantes a_1 y a_2 tal que $a_1(1,3,5) + a_2(2,4,6) = 0$. Por lo tanto, nosotros necesitamos $a_1 + 2a_2 = 0$, $3a_1 + 4a_2 = 0$ y $5a_1 + 6a_2 = 0$. Restando 3 veces la primera ecuación desde la segunda nos da $2a_2 = 0$ así que tenemos $a_1 = a_2 = 0$, lo que nos da finalmente que las filas de $\boldsymbol{X}$ son linealmente independientes y $\rho(\boldsymbol{X}) \geq 2$, por tanto $\rho(\boldsymbol{X}) = 2$.

Ejemplo 3:

Se tiene a $\boldsymbol{X} = \begin{bmatrix} 4 & 6 \\ 6 & 9 \end{bmatrix}$, una matriz de orden 2×2, por tanto el rango $\rho(\boldsymbol{X}) \leq 2$. Si $a_1(4,6) + a_2(6,9) = 0$ (por ejemplo haciendo $2a_1 + 3a_2 = 0$), entonces tenemos que $4a_1 + 6a_2 = 0$ y $6a_1 + 9a_2 = 0$, así podemos tener que $a_1 = 3$ y $a_2 = -2$, lo que lleva a que las columnas de $\boldsymbol{X}$ sean linealmente independientes y por lo tanto $\rho(\boldsymbol{X}) < 2$, pero $\rho(\boldsymbol{X}) \geq 1$, así concluimos que $\rho(\boldsymbol{X}) = 1$.

Ejemplo 4:

Se tiene a $\boldsymbol{X} = \begin{bmatrix} 1 & 2 & 3 & 2 \\ 4 & 5 & 6 & -1 \\ 5 & 7 & 9 & 1 \end{bmatrix}$, una matriz de orden 3×4, por tanto el rango $\rho(\boldsymbol{X}) \leq \min(3,4) = 3$. Mirando la matriz nos damos cuenta que la primera fila más la segunda fila es igual a la tercera fila, así que las filas no son linealmente independientes, por lo tanto $\rho(\boldsymbol{X}) < 3$.

- Howard, J. P. (2017). Computational Methods for Numerical Analysis with R. CRC Press.
- Banerjee, S., & Roy, A. (2014). Linear algebra and matrix analysis for statistics. Crc Pr