Álgebra Lineal Numérica Para Aprendizaje Estadístico

Notas de Estudio del Curso 2026

Miguel Trias

2026-04-29

Éste libro compila el material de estudio del curso Álgebra Lineal Para Aprendizaje Estadístico del año 2026. Está basado en los apuntes de clase del Profesor Diego Armentano y en las cátedras del Profesor Gilbert Strang.

Tabla de contenidos

Prefacio

Sobre este sitio

Este libro digital fue creado utilizando el ecosistema de **Código Abierto**:

- **Quarto**: El sistema de publicación científica y técnica de próxima generación.
- **RStudio IDE**: El entorno de desarrollo integrado para computación estadística.
- **Markdown**: El lenguaje de marcado ligero para la redacción de contenido.
- **GitHub Pages**: Para el alojamiento y despliegue automático del libro.

Todo el código fuente está disponible en el repositorio de Github.

Motivacion

Este libro digital surge como un esfuerzo personal para organizar y profundizar en los conceptos del curso **Álgebra Lineal para Aprendizaje Estadístico** de la **Facultad de Ciencias Económicas y Administración (UdelaR)**. El objetivo es recopilar toda la informacion disponible en los materiales del curso, enriquecidos con explicaciones y ejemplos asistidos por inteligencia artificial (**Gemini 1.5**).

Atribución Intelectual

Es fundamental aclarar que la autoría intelectual de los conceptos, teoremas y la estructura lógica presentados en este libro pertenece a las fuentes originales citadas. El núcleo teórico y la secuencia pedagógica provienen directamente de la obra del **Prof. Gilbert Strang (MIT)** (Strang, 2019), sus lecciones magistrales en *OpenCourseWare* (Strang, 2018), las notas de clase del **Prof. Diego Armentano** (Armentano, 2026) y los repartidos de ejercicios prácticos del curso (Hernández, 2026).

Mi rol en este proyecto se limita exclusivamente a la **curaduría, organización y edición de contenidos** para facilitar su estudio. Mi contribución personal se centra en la digitalización de los conocimientos, la creación de un formato interactivo mediante **Quarto**, y la generación de ejemplos aclaratorios asistidos por IA. Las contribuciones originales son ocasionales y tienen como único fin conectar las ideas de los autores mencionados o adaptar la notación para mayor claridad del lector.

⚠ Descargo de responsabilidad (Disclaimer)

Este material debe ser utilizado como una **guía de estudio complementaria** y no como una fuente oficial de verdad absoluta. Debido a que gran parte de la redacción y los ejemplos han sido procesados mediante IA:

- **No se ofrecen garantías sobre la exactitud del contenido.**
- Pueden existir errores técnicos, matemáticos o de transcripción que aún no han sido detectados.
- Se recomienda siempre contrastar los conceptos con la bibliografía oficial y las fuentes primarias citadas.

¡Colabora y Reporta!

Este es un proyecto abierto. Si encuentras una errata en una fórmula, una explicación confusa o quieres proponer una mejora en el código, no dudes en contribuir.

Puedes contribuir de las siguientes maneras:

1. **Reportando un error:** Abre un *Issue* en el repositorio de GitHub.
2. **Corrigiendo directamente:** Haz un *Fork* del proyecto y envía un *Pull Request* con tus correcciones.

Acceder al Repositorio en GitHub

Licencia y Uso

2. **No Comercial**: No puedes utilizar este material con fines comerciales.
3. **Compartir Igual**: Si remezclas o transformas este material, debes distribuir tus contribuciones bajo la misma licencia que el original.

1 Introducción

Esta introducción contiene una breve referencia sobre conocimientos previos, necesarios para entender el contenido del curso

1.1. Numeros Complejos

1.1.1. Propiedades

- Suma/Resta: sumamos las partes reales y las imaginarias
- Conjugado: Cambiamos el signo a la parte imaginaria
- Producto: Expresamos los numeros como $z = a + ib$ y aplicamos distributiva
- Cociente:
$$\frac{z_1}{z_2} = \frac{z_1\,\overline{z_2}}{|z_2|^2}, \quad \text{si } z_2 \neq 0$$

1.1.2. Interpretación geométrica

- **Suma:** suma de vectores en el plano
- **Multiplicación:** escala el módulo y **rota** el ángulo
- **Conjugado:** reflejo respecto al eje real
- **División:** divide módulos y **resta** ángulos

Observacion: El producto interno de complejos, (vector x vector = vector) no es generalizable a otros espacios. En $\mathbb{R}^3$ tenemos el producto vectorial usual, pero en $\mathbb{R}^2$ no hay un producto vectorial que sirva para definir un cuerpo.

1.1.3. Descomposición Polar de un Número Complejo

Cualquier número complejo z en forma binómica se escribe como $z = a + bi$, donde $a, b \in \mathbb{R}$ e i es la unidad imaginaria ($i^2 = -1$). La **descomposición polar** expresa este mismo punto utilizando coordenadas polares (r, θ).

- **El Módulo** r representa la magnitud o "longitud" del vector que une el origen con el punto z. Se obtiene mediante el teorema de Pitágoras aplicado a las componentes real e imaginaria: $r = |z| = \sqrt{a^2 + b^2}$
- **El Argumento** θ es el ángulo formado por el vector con el eje real positivo. Se define mediante la función arcotangente, teniendo en cuenta el cuadrante donde se ubica z: $\theta = \arg(z) = \arctan\left(\frac{b}{a}\right)$

1.1.3.1. Representación Matemática

Existen tres formas estrechamente vinculadas para expresar la descomposición polar:

1. **Forma Trigonométrica**
   Utilizando proyecciones sobre los ejes:

$$\begin{cases} a = r\cos\theta \\ b = r\sin\theta \end{cases} \implies z = r(\cos\theta + i\sin\theta)$$

2. **Forma Exponencial (Fórmula de Euler)**
   Basada en la identidad de Euler $e^{i\theta} = \cos\theta + i\sin\theta$, esta es la forma más compacta y potente para el análisis: $z = re^{i\theta}$
3. **Forma Polar (Notación de Fasores)**
   Común en ingeniería eléctrica y física: $z = r_{\angle\theta}$

1.1.3.2. Propiedades y Ventajas Operacionales

La descomposición polar simplifica drásticamente ciertas operaciones algebraicas en comparación con la forma binómica:

- **Multiplicación:**
  Se multiplican los módulos y se suman los ángulos.

$$z_1 z_2 = (r_1 r_2)e^{i(\theta_1+\theta_2)}$$

- **Potenciación (Fórmula de De Moivre):**

$$z^n = r^n(\cos(n\theta) + i\sin(n\theta))$$
$$(re^{i\theta})^n = r^n e^{in\theta}$$

- **Raíces:**
  Permite hallar las n raíces de un número complejo distribuyéndolas uniformemente en un círculo de radio $\sqrt[n]{r}$.

Observación:
De la formula de Euler:

$$e^{i\theta} = \cos\theta + i\sin\theta$$

Evaluando en $\theta = \pi$ surge

$$e^{i\pi} + 1 = 0$$

Observación:
$e^{i\pi}$ parametriza todos los puntos de $S^1 = z \in \mathbb{C} : |z| = 1$

1.2. Determinante de una Matriz Cuadrada

El determinante de una matriz cuadrada $A \in \mathbb{R}^{n\times n}$ es un escalar que representa el **factor de escala de volumen** de la transformación lineal asociada. Geométricamente, el volumen de un hiperparalelepípedo definido por las columnas de A es igual a $|\det(A)|$.

1.2.1. Propiedades Relevantes

1.2.1.1. Determinante del Producto de Matrices

$\det(AB) = \det(A)\det(B)$.

Idea de la prueba: Se basa en la unicidad del determinante como función multilineal alternada. Al definir $d(B) = \det(AB)$, se demuestra que $d(I) = \det(A)$, lo que fuerza la relación proporcional $d(B) = \det(A)\det(B)$.

1.2.1.2. Determinante como Producto de Autovalores

Si A tiene autovalores $\lambda_1, \dots, \lambda_n$: $\det(A) = \prod_{i=1}^{n} \lambda_i$.

Idea de la prueba: Evaluando el polinomio característico $p(\lambda) = \det(A - \lambda I)$ en $\lambda = 0$. La forma factorizada $(\lambda_1 - 0) \dots (\lambda_n - 0)$ coincide con el término constante del polinomio, que es $\det(A)$.

1.2.1.3. Determinante de Matrices Semejantes

Si $B = P^{-1}AP$, entonces $\det(B) = \det(A)$.

Idea de la prueba: Por la propiedad del producto, $\det(P^{-1}AP) = \det(P^{-1})\det(A)\det(P)$. Dado que $\det(P^{-1}) = 1/\det(P)$, los términos externos se anulan.

1.2.1.4. El Determinante en Matrices Ortogonales

Una propiedad fundamental de las matrices ortogonales ($Q^T Q = I$) es que su determinante está restringido a dos valores posibles: 1 **o** -1.

Prueba: Partiendo de la definición de matriz ortogonal: $Q^T Q = I$, aplicamos el determinante a ambos lados $\det(Q^T Q) = \det(I)$. Como $\det(I) = 1$ y el determinante del producto es el producto de los determinantes: $\det(Q^T)\det(Q) = 1$. Dado que el determinante de una transpuesta es igual al de la matriz original ($\det(Q^T) = \det(Q)$):

$$(\det(Q))^2 = 1 \implies \det(Q) = \pm 1$$

1.3. Independencia Lineal, Rango, Menores

El **rango** de una matriz se define como el número de columnas (o filas) linealmente independientes. Se dice que una matriz $A \in \mathbb{R}^{m\times n}$ de rango r, tiene **rango completo** si $r = \text{mín}(m, n)$.

Una herramienta analítica para determinar el rango de una matriz es el estudio de sus **menores**. Para una matriz de rango r, existe al menos un menor de tamaño $r \times r$ con determinante no nulo.

Si una matriz $A \in \mathbb{R}^{3\times 2}$ tiene rango 2, al menos uno de sus menores de 2×2 debe tener un determinante distinto de cero.

En una matriz genérica (con valores aleatorios), la probabilidad de que un menor sea singular es cero. Formalmente, en el espacio de matrices $\mathbb{R}^{3\times 2}$, el conjunto de matrices con rango menor a 2 (donde todos los menores de 2×2 son singulares) forma una **variedad algebraica**. Dado que las raíces de un polinomio no trivial definen un conjunto de medida cero, la probabilidad de que una matriz elegida aleatoriamente sea singular es nula.

2 Matrices de Rango 1

Si una matriz tiene rango 1, entonces todas sus columnas pueden obtenerse como combinaciones lineales de la primera (llamemosle u). Por lo tanto, esta matriz puede expresarse como el **producto exterior** de dos vectores u y $v^\top$, si elegimos las componentes adecuadas de v:

$$A = uv^\top$$

En esta estructura, el espacio columna es la recta generada por u y el espacio fila es la recta generada por v.

De forma general, cualquier matriz de rango r puede descomponerse en la suma de r matrices de rango 1:

$$A = \sum_{i=1}^{r} c_i r_i^\top$$

Esta perspectiva es esencial para entender factorizaciones avanzadas como SVD o la diagonalización, donde la información de la matriz se "fragmenta" en componentes de importancia decreciente.

Espacio Columna $C(A)$
Es el subespacio en el codominio $\mathbb{R}^m$ generado por las combinaciones lineales de las columnas de A. Para una matriz de rango 1, este espacio es simplemente la **recta** generada por el vector u.

Espacio Fila $C(A^T)$
Es el subespacio en el dominio $\mathbb{R}^n$ generado por las filas de A. En el caso de rango 1, es la **recta** generada por el vector v.

2.0.0.1. Ejemplo

Consideremos la matriz A de tamaño 2×2:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix}$$

Aquí, el vector $u = [1, 3]^\top$ define la dirección del espacio columna en $\mathbb{R}^2$, mientras que el vector $v = [1, 2]^\top$ define la dirección del espacio fila en $\mathbb{R}^2$.

2.1. Estructura Espectral

Las matrices de rango 1 son los bloques de construcción fundamentales del álgebra lineal numérica. Una matriz $A \in \mathbb{R}^{3\times 3}$ definida como el producto exterior de dos vectores no nulos, $A = uv^\top$, posee una geometría espectral muy simple pero poderosa. Este se divide en dos partes: el autovalor asociado a la imagen, o autovalor "vivo" y los autovalores asociados al núcleo o autovalores "nulos" (Strang, 2019, p. I.1; Armentano, 2026 clase 4).

El Autovalor "Vivo" λ_1
Si multiplicamos la matriz por el vector u, observamos:

$$Au = (uv^\top)u = u(v^\top u) = u\langle v, u\rangle = \langle v, u\rangle u$$

Esto revela que $\lambda_1 = v^T u = \langle v, u\rangle$ (el producto punto entre v y u) es el único autovalor que puede ser distinto de cero (Hernández, 2026, p. 3, ejercicio 12).

Los Autovalores Nulos λ_2, λ_3
Dado que el rango de la matriz es $r = 1$, la dimensión del núcleo es $n - r = 3 - 1 = 2$ (Armentano, 2026 clase 4). Por lo tanto, existen dos autovalores adicionales que deben ser cero: $\lambda_2 = 0, \quad \lambda_3 = 0$

2.2. Autovectores

La ubicación de los autovectores depende de los subespacios fundamentales de A (Strang, 2019, p. I.6):

Autovector para $\lambda_1 = v^T u$
Como se demostró arriba, el vector $x_1 = u$ es el autovector asociado al valor propio no nulo. Este vector define la dirección del espacio columna $C(A)$.

Autovectores para $\lambda = 0$
Los autovectores asociados al cero son todos los vectores que pertenecen al núcleo $N(A)$. Por definición, estos son los vectores perpendiculares al espacio fila (generado por v) (Armentano, 2026 clase 4):

$$Ax = (uv^\top)x = u(v^\top x) = 0 \iff v^\top x = 0$$

Para completar la base de autovectores en $\mathbb{R}^3$, simplemente elegimos dos vectores x_2, x_3 que sean linealmente independientes y perpendiculares a v

2.3. Resumen Espectral

Tipo	Valor Propio (λ)	Vector Propio	Ubicación Geométrica
Principal	$\lambda_1 = v^\top u = \langle u, v \rangle$	$x_1 = u$	$C(A)$
Nulo 1	$\lambda_2 = 0$	$x_2 \perp v$	$N(A)$
Nulo 2	$\lambda_3 = 0$	$x_3 \perp v$	$N(A)$

! Importante

La matriz es diagonalizable solo si $\lambda_1 \neq 0$ (es decir, $u \not\perp v$).
Si $v^\top u = 0$, la matriz tiene todos sus autovalores en cero y no posee una base completa de autovectores, convirtiéndose en una matriz defectiva (Strang, 2019, p. I.6).

2.4. El Polinomio Característico

La estructura espectral de una matriz de rango 1, $A = uv^\top$, se refleja de manera simplificada en su polinomio característico $\chi_A(\lambda) = \det(A - \lambda\mathbb{I})$ (Armentano, 2026 clase 7; Strang, 2019, p. I.6).

Para una matriz $n \times n$ de rango 1, el polinomio característico siempre toma la forma:

$$\chi_A(\lambda) = (-\lambda)^{n-1}(v^\top u - \lambda)$$

En el caso específico de una matriz 3×3, esto se reduce a:

$$\chi_A(\lambda) = -\lambda^2(\lambda - v^\top u) = -\lambda^3 + (v^\top u)\lambda^2$$

2.4.1. Observaciones

Multiplicidad Algebráica del Cero
El factor λ^{n-1} confirma que el valor propio $\lambda = 0$ tiene una multiplicidad algebraica de al menos $n - 1$. Esto es una consecuencia directa de que la matriz tiene un núcleo de dimensión $n - 1$ (Armentano, 2026 clase 4).

Relación con la Traza
El coeficiente del término λ^{n-1} (en este caso $v^T u$) corresponde a la **traza** de la matriz (Strang, 2018 Lec. 4 '144). En una matriz de rango 1, la traza es exactamente el producto interno de los vectores que la generan: $tr(uv^\top) = v^\top u$.

Determinante Nulo
Como el término constante del polinomio es el determinante y aquí es cero (para $n > 1$), se confirma que toda matriz de rango 1 es singular (Armentano, 2026 clase 4; Strang, 2019, p. I.6).

2.4.2. El Caso $\lambda = 0$ repetido

Si los vectores son ortogonales ($v^\top u = 0$), el polinomio se convierte en $\chi_A(\lambda) = (-\lambda)^n$. En este escenario, todos los autovalores son cero, pero la matriz solo posee $n-1$ autovectores independientes, lo que la hace no diagonalizable.

3 Producto de Matrices, Factorizacion CR

Este capítulo explora la dualidad en la interpretación de la multiplicación matriz-vector y la estructura fundamental de una matriz a través de sus columnas.

3.1. Multiplicación de Matrices por vector Ax

Sea $A = [a_1, a_2, \dots, a_n] \in \mathcal{M}^{m\times n}$ una matriz donde a_j representa su j-ésima columna. Sea $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ Entonces, el producto de la matriz A por el vector x es:

$$Ax = x_1a_1 + x_2a_2 + \cdots + x_na_n$$

Esta operación fundamental, puede interpretarse de dos maneras (Strang, 2018 Lec. 1):

Perspectiva de Filas:
Cada componente del vector resultante es el producto interno (tambien llamado producto punto) entre una fila de A y el vector x.

Perspectiva de Columnas:
El producto Ax es una **combinación lineal** de las columnas de A, donde los coeficientes son las componentes de x.

> Nota
>
> El vector resultante $b = Ax$ reside necesariamente en el subespacio generado por las columnas de la matriz A. De otra manera, el vector solución x, muestra como expresar el termino independiente b como combinación lineal de las columnas de A. Para algunos b, esto es imposible porque no estan en el espacio generado por las columnas de A.

La interpretación como combinacion lineal de las columnas de A es un pilar central del curso, ya que permite visualizar el producto como un movimiento dentro del subespacio generado por las columnas de la matriz.

Otra manera de visualizar el producto Ax en la perspectiva de columnas, es que, tomando un vector aleatorio x de $\mathbb{R}^n$ obtenemos un vector del espacio $C(A)$. A su vez, si pensamos a A como la matriz que representa a una cierta transformacion lineal $T : \mathbb{R}^m \to \mathbb{R}^n$ entonces $C(A)$ corresponde al subespacio $\mathrm{Im}(T)$

Es facil ver como consecuencia de lo anterior, que la dimension del espacio generado por A coincide con su rango. $\dim(C(A)) = \dim(\text{Im}(t)) = \text{rango}(A)$

Ejemplo
En este ejemplo, la primer igualdad corresponde a la interpretacion usual en el cálculo, donde cada entrada de la matriz resultante es el producto punto entre las filas de la matriz A y el vector x. La segunda igualdad, corresponde a la interpretacion de Ax como combinacion lineal de las columnas de A. (Strang, 2019, p. 1.1)

$$Ax = \begin{pmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 + 3x_2 \\ 2x_1 + 4x_2 \\ 3x_1 + 7x_2 \end{pmatrix} = x_1 \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 3 \\ 4 \\ 7 \end{pmatrix}$$

3.2. Teorema de Factorización $A = CR$

(Strang, 2019, p. 1.1) Sea $A \in \mathbb{R}^{m\times n}$ una matriz de rango r. Existen una matriz $C \in \mathbb{R}^{m\times r}$, y una matriz $R \in \mathbb{R}^{r\times n}$ tales que: [1]

$$A = CR$$

- La matriz $C \in \mathbb{R}^{m\times r}$ se construye seleccionando las primeras r columnas linealmente independientes de A.
- La matriz $R \in \mathbb{R}^{r\times n}$ contiene los coeficientes necesarios para reconstruir todas las columnas de A a partir de la base C. Notablemente, R se encuentra a menudo en forma escalonada reducida (o contiene un bloque identidad I_r).

Ejemplo:

(Strang, 2018 Lec. 1 '20)

$$A = \begin{pmatrix} 2 & 1 & 3 \\ 3 & 1 & 4 \\ 5 & 7 & 12 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 3 & 1 \\ 5 & 7 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

Observaciones:

- Las columnas de C forman una base de $C(A)$.
- Las filas de R forman una base de $C(A^T)$.

[1] Esta factorización permite la **reducción de dimensionalidad**. Si $r \ll n$ y $r \ll m$, la matriz puede almacenarse utilizando $r(m+n)$ coeficientes en lugar de mn. Esto es la base conceptual de técnicas de compresión y de aproximación de matrices de bajo rango en ciencia de datos.

3.3. Igualdad del Rango por Filas y Columnas

Un pilar del álgebra lineal es que $\dim(C(A)) = \dim(C(A^T)) = r$. La factorización $A = CR$ ofrece una prueba constructiva:

- Las columnas de C son una base de $C(A)$ por construcción ($dim = r$).
- De la igualdad $A = CR$, se observa que cada fila de A es una combinación lineal de las filas de R (los coeficientes de dicha combinación se encuentran en las filas de C). Como R tiene r filas independientes, el espacio fila tiene dimensión r, igual a la dimensión del espacio de columnas.

Observar: (Strang, 2018 Lec. 1 '29)

- Rango de columnas de A = # columnas de C = rango de filas de A = # filas de R = rango(A)
- Obtenemos las columnas de A como combinaciones lineales de las columnas de C utilizando como coeficientes las columnas de R
- Obtenemos las filas de A como combinaciones lineales de las filas de R utilizando como coeficientes las filas de C

3.4. Extensión a la Multiplicación de Matrices AB

La interpretación del producto Ax como una combinación lineal de columnas se extiende de manera natural al producto de dos matrices $A \in \mathbb{R}^{m\times n}$ y $B \in \mathbb{R}^{n\times p}$. Existen dos perspectivas fundamentales para descomponer esta operación.

3.4.1. Perspectiva de Columnas

Si denotamos las columnas de B como $b_1, b_2, \dots, b_p$, el producto AB puede visualizarse como una colección de productos matriz-vector:

$$AB = [Ab_1, Ab_2, \dots, Ab_p]$$

Bajo esta óptica, **cada columna de AB es una combinación lineal de las columnas de A**, donde los coeficientes de la j-ésima combinación provienen de la j-ésima columna de B. Esto confirma que el espacio columna del producto está contenido en el espacio columna de la matriz de la izquierda: $C(AB) \subseteq C(A)$.

3.4.2. Perspectiva de Suma de Productos Externos

Al multiplicar $A \in \mathbb{R}^{m \times n}$ por $B \in \mathbb{R}^{n \times p}$, cada columna k de A se encuentra exclusivamente con la fila k de B.

El producto de la columna k de A $(m \times 1)$ por la fila k de B $(1 \times p)$ genera una matriz completa de tamaño $m \times p$.

Este producto, denominado **producto exterior**, produce una matriz donde todas las columnas son múltiplos de la columna de A y todas las filas son múltiplos de la fila de B, lo que define una matriz de rango 1.

La matriz final AB es simplemente la suma de estos n bloques de construcción de rango 1 (Strang, 2018 Lec. 1, '7):

$$AB = \sum_{k=1}^{n} a_k b_{*k} = a_1 b_{*1} + a_2 b_{*2} + \cdots + a_n b_{*n}$$

donde:

- a_k es la k-ésima **columna** de A.
- b_{*k} es la k-ésima **fila** de B.

Cada término $a_k b_{*k}$ genera una matriz donde todas las columnas son múltiplos de a_k y todas las filas son múltiplos de b_{*k}.

Veamos por que el elemento $(AB)_{ij}$ de la matriz resultante es idéntico en ambos métodos.

La entrada (i, j) de la k-ésima matriz de la suma es exactamente $a_{ik} b_{kj}$. Al sumarlas todas, recuperamos la fórmula estándar: $\sum_{k=1}^{n} a_{ik} b_{kj}$.

En el método de productos internos (filas por columnas), tenemos:

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

En la suma de productos exteriores, el elemento (i, j) de la k-ésima matriz $a_k b_{*k}$ es precisamente $a_{ik} b_{kj}$. Al sumar las n matrices, obtenemos exactamente la misma sumatoria para cada entrada de la matriz final.

Esta visión de sumar piezas de rango 1 es un pilar de las grandes factorizaciones del curso: **Factorización** LU (Resta sucesivamente matrices de rango 1 para eliminar entradas) y **SVD y Descomposición Espectral** (Reconstruyen la matriz original sumando sus componentes más importantes [de mayor a menor valor singular/autovalor]).

Nota

Desde el punto de vista computacional, multiplicar una matriz $m \times n$ por una $n \times p$ requiere mnp multiplicaciones, independientemente del orden o la perspectiva utilizada.

4 Los Cuatro Subespacios Fundamentales

Toda matriz $A \in \mathbb{R}^{m\times n}$ actúa como una transformación lineal $T : V \to W$ que vincula subespacios vectoriales específicos en el dominio ($\mathbb{R}^n$) y en el codominio ($\mathbb{R}^m$). Esta acción se comprende a través de **cuatro subespacios fundamentales**, organizados en pares de complementos ortogonales. La comprensión de estos subespacios es fundamental para el análisis de solubilidad y estabilidad numérica. El objetivo de esta sección es comprender el "panorama" global de una matriz a través de estas cuatro piezas básicas que la componen.

4.1. El Espacio Columna $C(A)$

El **espacio columna** de una matriz $A \in \mathbb{R}^{m\times n}$, denotado como $C(A)$, es el subespacio de $\mathbb{R}^m$ formado por todas las combinaciones lineales posibles de sus columnas.

Definición:

$$C(A) = \{b \in \mathbb{R}^m : b = Ax \text{ para algún } x \in \mathbb{R}^n\}$$

Geométricamente, si las columnas de una matriz 3×2 son linealmente independientes, el espacio columna representa un plano en $\mathbb{R}^3$ que pasa por el origen. El sistema de ecuaciones $Ax = b$ tiene solucion si y solo si el vector b pertenece al espacio columna $C(A)$.

Si $b \notin C(A)$, el sistema se denomina **incompatible** o inconsistente. En tales casos, el álgebra lineal numérica busca la "mejor aproximación" mediante métodos como mínimos cuadrados, proyectando b sobre el subespacio $C(A)$.

4.2. El Espacio Fila $C(A^\top)$

Similarmente podemos definir el espacio Fila de A como el espacio generado por las filas de la matriz. Para no introducir otro simbolo, observamos que las filas de A son las columnas de su transpuesta, por tanto utilizamos la notacion $C(A^\top)$

4.3. Definiciones Formales

Espacio Columna $C(A)$ o Col(A) El subespacio de $\mathbb{R}^m$ generado por las combinaciones lineales de las columnas de A. Representa la **imagen** de la transformación lineal, es decir, todos los vectores b para los cuales el sistema $Ax = b$ tiene solucion.

Espacio Fila $C(A^\top)$ El subespacio de $\mathbb{R}^n$ generado por las combinaciones lineales de las filas de A. Reside en el dominio $\mathbb{R}^n$. Contiene las direcciones de los datos de entrada que la matriz realmente "ve" y transforma hacia la imagen.

Núcleo o Espacio Nulo $N(A)$ o Ker(A) El conjunto de todos los vectores $x \in \mathbb{R}^n$ tales que $Ax = 0$. Reside en el dominio $\mathbb{R}^n$. Estos vectores representan la información que se "pierde" o se colapsa al origen durante la transformación.

Núcleo Izquierdo $N(A^\top)$ El conjunto de todos los vectores $y \in \mathbb{R}^m$ tales que $A^\top y = 0$. Reside en el codominio $\mathbb{R}^m$. Contiene los vectores y tales que $A^\top y = 0$. Geométricamente, representa las combinaciones lineales de las filas de A que resultan en el vector nulo.

4.4. Teorema de las Dimensiones

(tambien llamado Teorema del Rango o Teorema fundamental del algebra)

Para cualquier matriz $A \in \mathbb{R}^{m\times n}$ de rango r que actúa como una transformación lineal $A : \mathbb{R}^n \to \mathbb{R}^m$, el Teorema de las Dimensiones establece que la dimensión del dominio se compone por la suma de las dimensiones del núcleo y de la imagen:

$$n = \dim N(A) + \dim C(A) \tag{4.1}$$

En términos prácticos, esto significa que el rango (r) más la nulidad $(n - r)$ siempre es igual al número de columnas de la matriz. $n = \dim N(A) + r$

4.5. Descomposición Ortogonal

Al resolver el sistema homogéneo $Ax = 0$, buscamos el núcleo o espacio nulo $N(A)$, compuesto por todos los vectores x que la transformación lineal A mapea al vector nulo.

Desde una perspectiva algebraica, esto equivale a encontrar los coeficientes de una combinación lineal de las columnas de A que resulte en cero; si existen vectores $x \neq 0$ que cumplen esto, las columnas son linealmente dependientes.

Desde una perspectiva geométrica, la ecuación $Ax = 0$ implica que el producto escalar de cada fila de A con el vector x es igual a cero $(r_i \cdot x = 0)$. Por lo tanto, el espacio nulo es el conjunto de vectores ortogonales al espacio de filas $C(A^\top)$, esto es, el **complemento ortogonal** (Strang, 2018 Lec. 2 '35)

$$N(A) = C(A^\top)^\perp$$

Esto implica que para todo $v \in N(A)$ y todo $w \in C(A^\top)$, se cumple $\langle v, w \rangle = 0$.

Otra forma de verlo, es que al multiplicar Ax, las componentes del vector resultante son los productos internos de cada fila $A_{(i)}$ de A con x $\langle A_{(i)}, x_i \rangle$, y estos tienen que dar 0. Esto implica que cualquier vector en el núcleo de A es perpendicular a todas las filas de la matriz.

Esto significa que el espacio de partida se puede descomponer como una suma directa de estos subespacios perpendiculares. Es decir, el dominio $\mathbb{R}^n$ admite una **descomposición ortogonal** dada por:

$$\mathbb{R}^n = N(A) \oplus C(A^\top)$$

De manera análoga, para el codominio se verifica:

$$N(A^\top) = C(A)^\perp$$

Por lo que:

$$\mathbb{R}^m = C(A) \oplus N(A^\top)$$

Utilizando estas propiedades de los subespacios fundamentales, podemos ver que Al aplicar la transformacion A a un vector x; si éste tiene una componente en el núcleo, esa parte será anulada por A, mientras que la componente en el espacio fila será mapeada biunívocamente hacia el espacio columna.

Por lo tanto, para conocer totalmente la transformacion A solo nos interesa conocer como transforma los vectores en $C(A^\top)$ porque los vectores de $N(A)$ siempre van a parar al 0.

4.5.1. Corolario: Igualdad del rango por filas y columnas

Al tomar dimensiones en ambos lados de la suma directa:

$$n = \dim N(A) + \dim C(A^\top)$$

Lo que a su vez, junto con la Ecuación 4.1 implica:

$$\dim C(A^\top) = \dim C(A)$$

Esta igualdad revela que toda la información linealmente independiente de una matriz se refleja de manera simétrica tanto en sus filas como en sus columnas. Mientras que las filas describen las restricciones sobre las soluciones del sistema ($Ax = 0$), las columnas describen la imagen o el alcance de la transformación. El rango r cuantifica estas “direcciones independientes” de manera unívoca para ambos espacios.

El rango por filas es idénticamente igual al rango por columnas ($\dim C(A) = \dim C(A^\top) = r$). Este hecho garantiza que la matriz A y su traspuesta $A^\top$ comparten la misma capacidad de generación de dimensiones independientes en sus respectivos subespacios.

4.5.2. Resumen

$A \in \mathcal{M}^{m\times n}$ de rango r, representa una transformacion lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ que puede caracterizarse a travez de los subspacios:

Subespacio	Símbolo	Dimensión	Espacio
Espacio Columna	$C(A)$	r	$\mathbb{R}^m$
Núcleo Izquierdo	$N(A^\top)$	$m - r$	$\mathbb{R}^m$
Espacio Fila	$C(A^\top)$	r	$\mathbb{R}^n$
Núcleo	$N(A)$	$n - r$	$\mathbb{R}^n$

4.6. El "Gran Mapa" del Álgebra Lineal

La acción de A puede visualizarse como una correspondencia biunívoca entre el espacio fila y el espacio columna (Armentano, 2026 Clase 4). Los vectores en $N(A)$ son mapeados al vector nulo en el codominio. Si restringimos el dominio de A únicamente al subespacio $C(A^\top)$, el operador resultante:

$$A : C(A^\top) \to C(A)$$

es un isomorfismo (es decir, es invertible sobre estos subespacios específicos).

Esto implica que para cada vector en la imagen $C(A)$, existe un único vector en el espacio fila $C(A^\top)$ que lo produce. Esta correspondencia es la base para entender la pseudo-inversa y la resolución de sistemas mediante SVD.

4.7. Restricciones y Codimensión

Para comprender la dimensión del núcleo $N(A)$, resulta útil visualizar las filas de la matriz $A \in \mathbb{R}^{m\times n}$ como una serie de restricciones geométricas aplicadas sobre el espacio ambiente $\mathbb{R}^n$.

Un vector x pertenece a $N(A)$ si y solo si es ortogonal a cada vector fila a_{i*} de la matriz A. Si denotamos el rango de la matriz como r, la existencia de r filas independientes impone r restricciones lineales sobre el dominio, lo que reduce la dimensión del conjunto solución a $n - r$.

Por esta razon es util hablar de **espacios de codimension** r, es decir, todo lo que le falta al espacio para abarcar la dimension del ambiente. Esta manera de expresarlo es útil porque no depende de la dimension del espacio de partida n.

4.7.1. La Fila como Hiperplano

Consideremos una única ecuación lineal $a_{1*}x = 0$, donde a_{1*} es la primera fila de la matriz. Geométricamente, esta ecuación define un subespacio de **codimensión 1**.

- Si estamos en $\mathbb{R}^n$, el conjunto de soluciones es un hiperplano de dimensión $n - 1$.
- Este hiperplano es, por definición, el subespacio ortogonal al vector fila a_{1*}.

4.7.2. Intersección y Reducción de Dimensión

Cuando añadimos una segunda fila a_{2*}, estamos buscando la intersección de dos subespacios.

- **Si las filas son linealmente independientes:** La intersección de estos dos hiperplanos es "transversal", lo que reduce la dimensión del espacio solución en exactamente uno más. Así, el espacio de soluciones pasa a tener dimensión $n - 2$.
- **Si las filas son dependientes:** Por ejemplo, si $a_{2*} = 2a_{1*}$, la segunda ecuación no aporta información nueva. La intersección de los subespacios es el mismo hiperplano original, y la dimensión no decrece.

4.7.3. Conclusión sobre la Dimensión del Núcleo

Este proceso se generaliza para las m filas de la matriz. Si la matriz tiene rango r (es decir, r filas linealmente independientes), la dimensión del núcleo se calcula como:

$$\dim(N(A)) = n - r$$

Cada "dirección independiente" capturada por el espacio fila $C(A^\top)$ consume una dimensión del dominio, dejando el resto para el núcleo. Esto refuerza la idea de que el núcleo y el espacio fila son complementarios: lo que una fila independiente "restringe", se pierde para el núcleo pero se gana para la descripción del sistema.

4.7.4. Ejemplo de Codimensión en $\mathbb{R}^3$

Si tenemos una matriz con $n = 3$ y una sola fila independiente ($r = 1$), el núcleo es un plano (dimensión $3 - 1 = 2$). Al añadir una segunda fila independiente ($r = 2$), el núcleo colapsa a una recta (dimensión $3 - 2 = 1$). Si añadimos una tercera fila independiente, el único vector perpendicular a las tres direcciones es el vector nulo (dimensión 0).

5 Ortogonalidad, Isometrías, Matrices de Stiefel, Proyectores

Las matrices ortogonales, representan las transformaciones más estables en el cálculo numérico. Estas preservan la geometría del espacio, lo que garantiza que los errores de redondeo no se amplifiquen durante las factorizaciónes.

5.1. Fundamentos de Ortogonalidad

5.1.1. Definición: Vectores Ortonormales

Dos vectores $x, y \in \mathbb{R}^n$ se definen como ortogonales si su producto escalar es nulo:

$$x^\top y = \sum_{i=1}^{n} x_i y_i = 0$$

Para el caso complejo ($\mathbb{C}^n$), la definición requiere el uso del conjugado traspuesto (producto hermitiano) para asegurar que la norma sea un número real no negativo:

$$x^H y = \overline{x}^\top y = \sum_{i=1}^{n} \overline{x_i} y_i = 0$$

5.1.2. Definición: Conjuntos y Bases Ortonormales

Un conjunto de vectores $\{q_1, q_2, \dots, q_k\}$ es **ortonormal** si cumple la condición de la delta de Kronecker:

$$q_i^\top q_j = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Decimos que un conjunto de vectores es una **base ortonormal** si es un conjunto ortogonal de vectores unitarios ($\|q_i\| = 1 \quad i = 1, 2, \dots, k$) que genera todo el espacio.

5.2. Representacion de Vectores en Bases no Ortonormales

Sea $\mathcal{B} = \{b_1, b_2, \dots, b_n\}$ una base **no ortogonal** de V. Para expresar un vector $v \in V$ como una combinación lineal de los elementos de dicha base, debemos encontrar un conjunto de escalares $\{c_i\}$ tales que:

$$v = \sum_{i=1}^{n} c_i b_i$$

Este planteamiento requiere resolver el sistema lineal $Bc = v$, donde B es la matriz cuyas columnas son los vectores de la base. Esta operación conlleva varias dificultades desde una perspectiva computacional. [1]

5.3. Representación de Vectores en Bases Ortonormales

Cualquier vector $v \in \mathbb{R}^n$ puede expresarse de manera única como una combinación lineal de las columnas $q_1, \dots, q_n$ de una matriz ortogonal Q:

$$v = c_1 q_1 + c_2 q_2 + \dots + c_n q_n$$

Multiplicando la expresión de v por $q_1^\top$ a la izquierda:

$$q_1^\top v = c_1(q_1^\top q_1) + c_2(q_1^\top q_2) + \dots + c_n(q_1^\top q_n)$$

Por la propiedad de ortonormalidad, $q_1^\top q_1 = 1$ y $q_1^\top q_j = 0$ para todo $j \neq 1$, resultando en $c_1 = q_1^\top v$.

Conclusión

Sea $Q = \{q_1, q_2, \dots, q_n\}$ una base ortonormal de $\mathbb{R}^n$.
$\forall v \in \mathbb{R}^n : v = c_1 q_1 + c_2 q_2 + \dots + c_n q_n$
Los coeficientes c_i se obtienen mediante productos internos individuales: $c_i = q_i^\top v$
En notación matricial, el vector de coeficientes completo se obtiene aplicando la matriz

[1]

- **Complejidad Algorítmica:** La resolución mediante métodos directos (como la eliminación gaussiana o la descomposición LU) implica un costo de $O(n^3)$ operaciones de punto flotante.
- **Interdependencia de Coeficientes:** El cálculo de cualquier coeficiente c_i está acoplado al resto del sistema, lo que impide una resolución parcial o asíncrona de las coordenadas. A diferencia de las bases arbitrarias, donde el cálculo de las coordenadas exige resolver un sistema lineal mediante eliminación ($O(n^3)$), en una base ortonormal los coeficientes se obtienen mediante productos internos individuales ($O(n)$):

$Q^\top$ al vector v:

$$c = Q^\top v$$

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} - & q_1^\top & - \\ \vdots & \vdots & \vdots \\ - & q_n^\top & - \end{bmatrix} v$$

5.4. Matrices de Stiefel

Una matriz Q de tamaño $n \times r$ se denomina **Matriz de Stiefel** si sus columnas son vectores ortonormales en $\mathbb{R}^n$.

Una consecuencia directa de esta definición es la identidad fundamental:

$$Q^\top Q = \mathbb{I}_r$$

donde $\mathbb{I}_r$ es la matriz identidad de tamaño $r \times r$.

$$Q^\top Q = \begin{bmatrix} - & q_1^\top & - \\ \vdots & \vdots & \vdots \\ - & q_r^\top & - \end{bmatrix} \begin{bmatrix} | & \cdots & | \\ q_1 & \cdots & q_r \\ | & \cdots & | \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & \ddots & 0 \\ 0 & 0 & \dots & 1 \end{bmatrix} = \mathbb{I}_r$$

! Importante

Si $n > r$, la matriz $QQ^\top$ no es la identidad, sino un operador de proyección sobre el subespacio generado por las columnas de Q.

5.4.1. Teorema: Preservación de la Norma (Propiedad Isométrica)

La multiplicación por una matriz de Stiefel preserva la norma euclídea de cualquier vector $x \in \mathbb{R}^r$. (Armentano, 2026 clase 4; Strang, 2018 Lec. 3 '6)

Demostración:
Sea $x \in \mathbb{R}^r$. La norma al cuadrado del vector transformado Qx es:

$$\|Qx\|^2 = (Qx)^\top (Qx) = x^\top (Q^\top Q) x = x^\top I_r x = x^\top x = \|x\|^2$$

Por lo tanto, $\|Qx\| = \|x\|$. Geométricamente, esto significa que Q actúa como una rotación o reflexión, manteniendo intactas las distancias y los ángulos originales del dominio.

5.5. El Operador de Proyección Ortogonal $QQ^\top$

Es imperativo distinguir entre las dos combinaciones posibles de una matriz de Stiefel $Q \in \mathbb{R}^{n\times r}$:

Matriz Identidad $Q^\top Q = \mathbb{I}_r$ Puede visualizarse como el producto interno de filas por columnas, donde cada entrada de la matriz resultante es el producto punto de una fila de Q por una columna. Este producto interno produce la identidad de tamaño r, confirmando su ortonormalidad.

Proyección Ortogonal $QQ^\top = P$ En este caso, ayuda ver a Q como un conjunto de columnas $q_1, q_2, \dots, q_r$, y a $QQ^\top$ como el producto de columnas por filas lo que convierte cada término de $q_i q_i^\top$ en un producto exterior. A diferencia del producto punto (que da un número), el producto exterior de un vector por sí mismo genera una matriz de rango 1 de tamaño $n \times n$.

$$P = q_1 q_1^\top + q_2 q_2^\top + \cdots + q_r q_r^\top$$

Si $r < n$, esta matriz es singular, de rango r y actúa como un **proyector ortogonal** sobre el espacio columna $C(Q)$.

Tip

$Q^\top Q = \mathbb{I}_r$	Identidad de tamaño r
$QQ^\top = P \implies QQ^\top v = Pv = v_Q$	Proyección Ortogonal sobre$C(Q)$

5.5.1. Prueba Geométrica y Algebraica

Para que un operador lineal sea una proyección ortogonal, debe satisfacer dos condiciones estructurales que garantizan la estabilidad y la geometría del mapeo.

5.5.1.1. Idempotencia

Aplicar la proyección sucesivamente no debe alterar el vector tras la primera aplicación. $P^2 = (QQ^\top)(QQ^\top) = Q(Q^\top Q)Q^\top = Q(I_r)Q^\top = QQ^\top = P$ Entonces, $P^2 = P$.

Si $v \in C(Q)$, entonces existe un x tal que $v = Qx$. Al aplicar el proyector: $Pv = P(Qx) = (QQ^\top)Qx = Q(Q^\top Q)x = Qx = v$. Por lo tanto, P deja invariantes a los vectores que ya pertenecen a su imagen.

5.5.1.2. Simetría

Si estamos proyectando un vector x sobre P, su proyección es Px y el error es $x - Px$. Para que esta proyección sea ortogonal, el error de la proyección debe ser perpendicular al espacio donde se proyecta, es decir, que el error debe ser perpendicular a cualquier vector Py en la imagen de P. Por lo tanto debe cumplirse

$$(Py)^\top(x - Px) = 0, \forall y$$

Desarrollando esta formula, $0 = y^\top P^\top(x - Px) = y^\top(P^\top x - P^\top Px) = y^\top(P^\top - P^\top P)x$. Para que esto sea cierto para todo x e y, debe cumplirse $P^\top = P^\top P$.

Si además la matriz P fuese simétrica ($P = P^\top$), entonces tendriamos $P^\top = P^\top P = PP = P$ Por lo tanto, la simetria de P garantiza que la proyección sea ortogonal y no oblicua. Verifiquemos entonces que P es simetrica: $P^\top = (QQ^\top)^\top = (Q^\top)^\top Q^\top = QQ^\top = P$.

5.5.2. Caracterización Geométrica

El operador P descompone cualquier vector $w \in \mathbb{R}^n$ en una componente dentro del subespacio y un error ortogonal: $w = Pw + (\mathbb{I} - P)w$.

Para cualquier vector w, el "vector de error" definido como $e = w - Pw$ es perpendicular a toda combinación lineal de las columnas de Q, es decir $w - Pw \perp C(Q)$.

Un vector es perpendicular a $C(Q)$ si su producto interno con cada columna de Q es nulo, lo cual equivale a decir que $Q^\top e = 0$.

$Q^\top e = Q^\top(w - Pw) = Q^\top(w - QQ^\top w) = Q^\top w - (Q^\top Q)Q^\top w = 0$

5.5.3. Aplicación: La Proyección como Mejor Aproximación

La proyección ortogonal es la piedra angular del método de **Mínimos Cuadrados**. Cuando un sistema $Ax = b$ es incompatible, es decir, $b \notin C(A)$, la solución numérica óptima no busca resolver la igualdad, sino minimizar el tamaño del error: $\|b - Ax\|$. En otras palabras, buscamos el vector $\hat{x}$ en $C(A)$ más cercano a b en términos de distancia euclídea.

Si las columnas de A se transforman en una base ortonormal Q, la "mejor aproximación" de b en el subespacio es simplemente su proyección ortogonal $Pb = QQ^\top b$.

Pero estamos interesados en encontrar una solucion para cualquier A. Como vimos antes, para que el error sea mínimo, el vector de error $(b - A\hat{x})$ debe ser perpendicular a todas las columnas de A. En lenguaje matricial, esto significa que el producto de $A^\top$ por el error debe ser cero:

$$A^\top(b - A\hat{x}) = 0$$

Si distribuimos el producto, obtenemos las famosas **Ecuaciones Normales**:

$$A^\top b - A^\top A\hat{x} = 0 \tag{5.1}$$

$$A^\top A\hat{x} = A^\top b \tag{5.2}$$

Ahora, para encontrar $\hat{x}$, multiplicamos por la inversa de $(A^\top A)$ (asumiendo que las columnas de A son linealmente independientes):

$$\hat{x} = (A^\top A)^{-1}A^\top b$$

Sabemos que la proyección de b sobre el espacio de A es, por definición, Pb. Pero también sabemos que esa proyección es el resultado de aplicar la matriz A a nuestra mejor aproximación $\hat{x}$:

$$Pb = A\hat{x}$$

Sustituimos el valor de $\hat{x}$ que acabamos de despejar:

$$Pb = A\left((A^\top A)^{-1}A^\top b\right) = A(A^\top A)^{-1}A^\top b$$

Si quitamos la b de ambos lados, nos queda la estructura de la matriz de proyección:

$$P = A(A^\top A)^{-1}A^\top$$

6 Matrices de Householder

Una **matriz de Householder** H es un tipo especial de matriz ortogonal y simétrica que se utiliza para reflejar vectores respecto a un hiperplano.

La expresión $vv^\top$ es un proyector sobre la dirección de v. Al calcular $(\mathbb{I} - vv^\top)u$, estamos restando la componente de u en la dirección de v, lo que "colapsa" el vector u sobre el plano (proyección).

Si en cambio restamos la componente de u en la dirección de v **dos veces** $(\mathbb{I} - 2vv^\top)$, obligamos al vector a cruzar el plano y quedar exactamente a la misma distancia en el lado opuesto. A esta transformacion se le llama **Matriz de Householder**.

6.0.1. Definición Formal

Dado un vector $v \in \mathbb{R}^n$ unitario ($\|v\| = 1$), la matriz de Householder se define como:

$$H = \mathbb{I} - 2vv^\top$$

La matriz H actúa como un operador de reflexión:

- Si un vector x es ortogonal a v, entonces $Hx = x$ (el vector reside en el hiperplano de reflexión).
- Si un vector es paralelo a v, entonces $Hv = -v$ (el vector se refleja completamente).

6.0.2. Propiedades Estructurales

La potencia de Householder en el cálculo numérico reside en sus propiedades estructurales:

Simetría La matriz es idéntica a su traspuesta. $H^\top = (I - 2vv^\top)^\top = \mathbb{I}^\top - 2(v^\top)^\top v^\top = \mathbb{I} - 2vv^\top = H$

Isometría Como es ortogonal, preserva la norma del vector original y nunca amplifica el error numérico. $\|Hu\| = \|u\|$.

Involución Aplicar la reflexión dos veces regresa el espacio a su estado original: $H^2 = (\mathbb{I} - 2vv^\top)(\mathbb{I} - 2vv^\top) = \mathbb{I} - 4vv^\top + 4v(v^\top v)v^\top$ Como $v^\top v = 1$, la expresión se simplifica a $I - 4vv^\top + 4vv^\top = \mathbb{I}$. Por lo tanto: $H^2 = \mathbb{I}$.

Y uniendo con la simetría obtenemos tambien, $H^{-1} = H^\top$ y $H^2 = \mathbb{I}$.

6.0.3. Utilidad en Factorizaciones

En la práctica, utilizamos Householder para “limpiar” columnas de una matriz. Si queremos que un vector x se convierta en un vector con ceros en todas las posiciones excepto la primera (proyectarlo sobre el eje e_1), diseñamos un espejo que esté justo a mitad de camino entre x y el eje deseado. La “limpieza” de una columna mediante una reflexión de Householder consiste en diseñar un hiperplano (espejo) que sea el bisector perpendicular del camino entre el vector original x y el objetivo deseado z (Armentano, 2026 clase 6; Strang, 2018 Lec. 3, ’70).

Generalmente se busca que toda la norma del vector x se concentre en la primera componente para crear ceros debajo. Por tanto, el objetivo es $z = \|x\|e_1$. El vector perpendicular al espejo debe ser la línea que conecta ambos puntos: $v = x - z$. Este vector v apunta directamente desde el origen del vector hacia su imagen reflejada. Para asegurar que la matriz sea ortogonal, se utiliza el vector unitario $u = \frac{v}{\|v\|}$. La matriz resultante $H = I - 2uu^\top$ funciona restando exactamente el doble de la proyección de x sobre el vector normal u. Matemáticamente:

$$Hx = (I - 2uu^\top)x = x - 2u(u^\top x)$$

Al aplicar Hx, la matriz resta exactamente el doble de la proyección de x sobre el vector normal u. Esto cancela toda la parte de x que “sobra” fuera del eje e y lo empuja a aterrizar exactamente en el objetivo z, dejando la columna “limpia” con ceros debajo de la primera posición

7 Autovalores, Autovectores, $A = X\Lambda X^{-1}$

El estudio de los autovalores y autovectores permite entender la acción de una matriz A no a través de sus entradas individuales, sino a través de sus direcciones invariantes. Mientras que la mayoría de los vectores cambian de dirección al ser multiplicados por A, los autovectores permanecen en la misma línea, experimentando únicamente un escalamiento.

7.1. La Ecuación Fundamental

Sea $A \in \mathbb{R}^{n\times n}$ una matriz cuadrada. Un vector no nulo x es un **autovector** (o eigenvector) de A si existe un escalar λ, denominado **autovalor** (o eigenvalor), tal que:

$$Ax = \lambda x \tag{7.1}$$

Geométricamente, la aplicación de A sobre x no altera su dirección, solo su magnitud por un factor λ.

Esta definición es exclusiva para matrices cuadradas porque el transformado de x, es decir, Ax está en el mismo subespacio vectorial que x.

> **Nota**
>
> La definicion de autovector requiere que $x \neq \vec{0}$. Sin embargo, el autovalor 0 si es permitido. De hecho, cualquier vector en el núcleo es un autovector con $\lambda = 0$.

> **Tip**
>
> Si la matriz es invertible y $\lambda \neq 0$, se verifica que x también es autovector de A^{-1} con autovalor $1/\lambda$. Observar que en este caso, λ necesariamente tiene que ser distinto de 0 para que esa igualdad tenga sentido, pero el caso $\lambda = 0$ tiene sentido igualmente. $\lambda = 0$ significa que $Ax = 0$ es decir, que $x \in N(A)$, lo cual a su vez implica que el rango de A es menor que n y que A no es invertible, por lo que la expresion $A^{-1}x = \frac{1}{\lambda}x$ no es valida tampoco del lado izquierdo.

7.2. Motivación: Convergencia de Algorítmos Numéricos

Una de las mayores utilidades de los autovectores es el cálculo de potencias de matrices. Si aplicamos A sucesivamente k veces sobre un autovector x:

$$A^k x = \lambda^k x$$

Esto implica que el comportamiento asintótico de A^k está gobernado por la magnitud de sus autovalores:

- Si $|\lambda| > 1$, la componente en esa dirección crece exponencialmente.
- Si $|\lambda| < 1$, la componente decae o se atenúa hacia cero.

Para la mayoría de los vectores, la multiplicación por una matriz A cambia tanto su longitud como su dirección. Sin embargo, los **autovectores** son direcciones especiales que permanecen invariantes: el producto Ax se mantiene en la misma línea que x, escalado únicamente por el **autovalor** λ ($Ax = \lambda x$).

Esta propiedad se vuelve extremadamente potente cuando calculamos potencias de la matriz. Si aplicamos A sucesivamente k veces sobre un autovector, el resultado es simplemente $A^k x = \lambda^k x$.

La verdadera utilidad aparece cuando una matriz $n \times n$ posee n autovectores linealmente independientes $\{x_1, x_2, \dots, x_n\}$. En este escenario, estos vectores forman una base completa del espacio $\mathbb{R}^n$ (Strang, 2018 Lec. 4 '10).

Cualquier vector arbitrario v puede expresarse como una combinación lineal de estos autovectores:

$$v = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n$$

donde los coeficientes c_i representan las coordenadas de v en esta "base natural" de la matriz.

Calcular $A^k v$ mediante multiplicaciones matriciales directas es computacionalmente costoso y numéricamente inestable. Sin embargo, al utilizar la base de autovectores, el cálculo se vuelve inmediato. Gracias a la linealidad, la acción de A^k se distribuye sobre cada componente:

$$A^k v = c_1(\lambda_1^k x_1) + c_2(\lambda_2^k x_2) + \cdots + c_n(\lambda_n^k x_n)$$

Bajo esta óptica:

1. **Descomponemos v:** Encontramos cuánta "magnitud" tiene v en cada dirección invariante.
2. **Evolucionamos el sistema:** Cada dirección evoluciona de forma independiente según la magnitud de su autovalor λ_i^k.
3. **Resultado Asintótico:** Si $|\lambda_i| < 1$, esa componente desaparecerá con el tiempo; si $|\lambda_i| > 1$, esa dirección dominará el crecimiento del sistema.

Esta perspectiva permite tratar a las matrices diagonalizables como si fueran simples matrices diagonales, facilitando el análisis de estabilidad y convergencia en algoritmos numéricos.

7.3. El Polinomio Característico

Para hallar los autovalores, reescribimos la Ecuación 7.1 como un sistema homogéneo singular $(A - \lambda\mathbb{I})x = 0$. Para que exista una solución no trivial, la matriz $(A - \lambda\mathbb{I})$ debe ser singular, lo que conduce al **polinomio característico** y el estudio de sus raices:

$$\chi_A(\lambda) = \det(A - \lambda\mathbb{I}) = 0$$

Las raíces de este polinomio de grado n constituyen el **espectro** de la matriz.

7.4. Espectro de la Traspuesta $A^\top$

Toda matriz cuadrada A comparte exactamente los mismos autovalores con su traspuesta $A^\top$, porque ambas generan el mismo polinomio caracteristico (Strang, 2019, p. I.6; Strang, 2018 Lec. 4 '131; Armentano, 2026 clase 7).

Sin embargo, aunque los autovalores son los mismos, los **autovectores** suelen ser diferentes. Los autovectores de $A^\top$ se conocen como **autovectores izquierdos** de A, ya que satisfacen $y^\top A = \lambda y^\top$. En el "Gran Mapa", esto significa que mientras los autovectores de A definen direcciones invariantes en el dominio, los de $A^\top$ definen direcciones invariantes respecto a las filas (Armentano, 2026 clase 4).

7.5. Espectro de Matrices Semejantes

Dos matrices A y B son **semejantes** si existe una matriz invertible X tal que $A = XBX^{-1}$.

Matrices semejantes comparten el mismo espectro de autovalores (pero no comparten autovectores). $Av = \lambda v \Rightarrow X^{-1}BXv = \lambda v \Rightarrow BXv = X\lambda v \Rightarrow B(Xv) = \lambda(Xv)$. Esta última igualdad implica que Xv es vector propio de B con valor propio λ.

Prueba Sean dos matrices A y B semejantes ($\exists M/B = M^{-1}AM$). La semejanza preserva el polinomio característico, incluso si los autovalores λ residen en el plano complejo $\mathbb{C}$ (Strang, 2018 Lec. 4 '125):

$$\det(B - \lambda\mathbb{I}) = \det(M^{-1}AM - \lambda M^{-1}M) = \det(M^{-1}(A - \lambda\mathbb{I})M)$$

Utilizando la propiedad distributiva del determinante ($\det(ABC) = \det(A)\det(B)\det(C)$):

$$\det(M^{-1})\det(A - \lambda\mathbb{I})\det(M) = \frac{1}{\det(M)}\det(A - \lambda\mathbb{I})\det(M) = \det(A - \lambda\mathbb{I})$$

Corolario El producto de matrices AB tiene los mismos valores propios (no nulos) que BA. Para probarlo usando la observacion anterior, deberiamos probar que AB y BA son semejantes. Esto equivale a buscar una matriz X tal que $X(AB)X^{-1} = BA$. Esta igualdad se satisface simplemente tomando $X = B$, ya que $B(AB)B^{-1} = BA(BB^{-1}) = BA\mathbb{I} = BA$

7.6. Descomposicion $X\Lambda X^{-1}$

Si una matriz A posee n autovectores linealmente independientes, estos pueden agruparse como columnas de una matriz X. La relación $AX = X\Lambda$ permite la **factorización por diagonalización** (Strang, 2018 Lec. 4 ’36):

$$A = X\Lambda X^{-1}$$

donde $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ es una matriz diagonal con los autovalores en su diagonal principal (Strang, 2018 Lec 4 ’13).

Bajo esta perspectiva, multiplicar por A equivale a:

1. Cambiar a la base de autovectores (X^{-1}).
2. Escalar cada componente por su respectivo λ_i (Λ).
3. Regresar a la base original (X).

7.7. Vínculo entre Espectro, Determinante y Traza

7.7.1. Determinante

El producto de los autovalores es igual al determinante de la matriz:

$$\det(A) = \prod \lambda_i$$

La forma más sencilla de demostrar que el determinante es el producto de los autovalores es evaluando el polinomio característico en un punto específico (Hernández, 2026 Practico 3 Ej 10). Por el Teorema Fundamental del Álgebra, el polinomio se puede expresar en términos de sus raíces (autovalores):

$$\chi_A(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \dots (\lambda_n - \lambda)$$

Si evaluamos la expresión anterior en $\lambda = 0$ obtenemos:

$$\chi_A(0) = (\lambda_1 - 0)(\lambda_2 - 0) \dots (\lambda_n - 0) = \lambda_1 \lambda_2 \dots \lambda_n$$

Por otro lado, por la definición original en $\lambda = 0$:

$$\chi_A(0) = \det(A - 0\mathbb{I}) = \det(A)$$

ℹ Espectro de Matrices Singulares

Como $\det(A) = \prod_{i=1}^{n} \lambda_i$, esta propiedad implica que una matriz es singular ($\det(A) = 0$) si y solo si al menos uno de sus autovalores es cero (Strang, 2019, p. I.6).

7.7.2. Traza

La suma de las entradas de la diagonal principal es igual a la suma de los autovalores:

$$\mathrm{tr}(A) = \sum \lambda_i$$

La relación con la traza surge al comparar los coeficientes del polinomio característico en sus dos formas de expansión (Strang, 2018 - Lec. 4, '144).

- **En la forma factorizada**: Al expandir el producto $(\lambda_1 - \lambda)(\lambda_2 - \lambda) \dots (\lambda_n - \lambda)$, el coeficiente del término $(-\lambda)^{n-1}$ es exactamente la suma de las raíces: $\sum \lambda_i$.
- **En el determinante**: Al expandir $\det(A - \lambda\mathbb{I})$ mediante la fórmula de Leibniz o cofactores, el único término que contiene $(-\lambda)^{n-1}$ proviene del producto de los elementos de la diagonal principal:
$$(a_{11} - \lambda)(a_{22} - \lambda) \dots (a_{nn} - \lambda)$$
En este producto, el coeficiente de $(-\lambda)^{n-1}$ es la suma de los elementos de la diagonal: $a_{11} + a_{22} + \cdots + a_{nn}$ (Strang, 2019, p. I.6).

Dado que ambos coeficientes deben ser iguales por pertenecer al mismo polinomio, se cumple que $\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i$.

7.7.3. Verificación en Matrices Diagonalizables

Si $A = X\Lambda X^{-1}$, podemos usar las propiedades de invarianza por semejanza (Strang, 2019, p. I.6):

- $\det(A) = \det(X)\det(\Lambda)\det(X)^{-1} = \det(\Lambda) = \prod \lambda_i$.
- $\mathrm{tr}(A) = \mathrm{tr}(X\Lambda X^{-1}) = \mathrm{tr}(X^{-1}X\Lambda) = \mathrm{tr}(\Lambda) = \sum \lambda_i$.

Estructura Espectral en Matrices Aleatorias

El estudio de matrices con coeficientes aleatorios permite entender cuán "comunes" son ciertas propiedades estructurales en el álgebra lineal. Si generamos una matriz $A \in \mathbb{R}^{n\times n}$ con entradas distribuidas continuamente (por ejemplo, mediante una distribución normal), podemos predecir su comportamiento espectral.

Existencia de Autovectores en el Plano Complejo

Desde un punto de vista algebraico, la probabilidad de que una matriz cuadrada tenga autovectores es del 100 % siempre que consideremos el campo de los números complejos $\mathbb{C}$ (Strang, 2019, p. I.6). Esto se debe a que el polinomio característico $\chi_A(\lambda) = \det(A - \lambda \mathbb{I})$ es de grado n y, por el Teorema Fundamental del Álgebra, posee exactamente n raíces (Strang, 2019, p. I.6).

Sin embargo, la existencia de autovectores en el campo de los reales $\mathbb{R}$ depende de la dimensión:

- n **impar:** La matriz siempre tendrá al menos un autovalor real (y por ende un autovector real), ya que el polinomio de grado impar necesariamente cruza el eje real (Strang, 2019, p. I.6).
- n **par:** Existe una probabilidad no nula de que todos los autovalores sean complejos conjugados, lo que implica la ausencia de autovectores reales (Armentano, 2026 clase 7).

Diagonalizabilidad casi segura

Para una matriz aleatoria, la probabilidad de que el polinomio característico tenga raíces múltiples es cero (Armentano, 2026 clase 7). Por lo tanto, si los autovalores son reales, la matriz es **diagonalizable casi con seguridad** (Armentano, 2026 clase 7). Esto tiene implicaciones fundamentales para la estructura de la matriz:

- **Autovalores distintos:** Casi con seguridad, la matriz tendrá n autovalores diferentes (Armentano, 2026 clase 7).
- **Independencia Lineal:** Una matriz con n autovalores distintos garantiza que sus autovectores asociados son linealmente independientes (Strang, 2019, p. I.6).
- **Conclusión:** Una matriz aleatoria es **diagonalizable casi con seguridad** (con probabilidad 1), permitiendo siempre la factorización $A = X\Lambda X^{-1}$ (Armentano, 2026 clase 7).

El Determinante y la Invertibilidad

La probabilidad de que una matriz aleatoria sea singular es cero (Armentano, 2026 clase 7). En términos espectrales, esto significa que el valor $\lambda = 0$ casi nunca es una raíz del polinomio característico $\det(A - \lambda \mathbb{I}) = 0$ (Strang, 2019, p. I.6).

Esta propiedad asegura que los algoritmos numéricos que dependen de la invertibilidad o del cálculo de potencias A^k (con k negativo o positivo) sean estables frente a perturbaciones aleatorias (Armentano, 2026 clase 7).

Resumen de Probabilidades Espectrales

Propiedad	Probabilidad	Razón
Diagonalizable	**1**	Los autovalores son casi siempre distintos (Armentano, 2026 clase 7).
Invertible $\det(A) \neq 0$	**1**	El conjunto de matrices singulares tiene medida cero (Armentano, 2026 clase 7).
Autovectores en $\mathbb{R}^n$ (n par)	$<$ **1**	Posibilidad de autovalores complejos conjugados (Strang, 2019, p. I.6).
Autovectores en $\mathbb{C}^n$	**1**	Teorema Fundamental del Álgebra (Strang, 2019, p. I.6).

La Probabilidad de Autovalores Reales

Si generamos una matriz $A \in \mathbb{R}^{n \times n}$ con entradas aleatorias (distribuidas de forma continua), la probabilidad de que sus autovalores sean reales depende fundamentalmente de la dimensión n y de las raíces de su polinomio característico.

Perspectiva del Polinomio Característico

Los autovalores son las raíces de $\chi_A(\lambda) = \det(A - \lambda \mathbb{I}) = 0$.

- **Dimensión Impar (n impar):** La probabilidad de tener al menos un autovalor real es **1**. Esto ocurre porque cualquier polinomio de grado impar con coeficientes reales tiende a $+\infty$ y $-\infty$ en sus extremos, lo que garantiza que cruce el eje real al menos una vez (Strang, 2019, p. I.6).

- **Dimensión Par (n par):** La probabilidad es **menor a 1**. Existe una probabilidad no nula de que todas las raíces sean pares de complejos conjugados, lo que resultaría en una matriz sin autovalores reales (Strang, 2019, p. I.6).

Perspectiva del Determinante y la Invertibilidad

El determinante es el producto de los autovalores: $\det(A) = \lambda_1 \lambda_2 \ldots \lambda_n$. Para matrices aleatorias, la probabilidad de que $\det(A) = 0$ es cero (Armentano, 2026 clase 7). Esto implica que la probabilidad de que el cero sea un autovalor es nula, asegurando que los autovalores reales existentes sean casi seguramente no nulos (Strang, 2019, p. I.6).

Ejemplos Conceptuales

Matrices de Rotación:
Una matriz de rotación pura en $\mathbb{R}^2$ tiene la forma: $Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. Sus autovalores son complejos ($e^{\pm i\theta}$), ilustrando por qué en dimensiones pares la probabilidad de autovalores reales no es la unidad (Strang, 2019, p. I.6).

Matrices Simétricas:
Si restringimos la matriz aleatoria para que sea simétrica ($A = A^T$), la probabilidad de tener todos sus autovalores reales es **1**, independientemente de n (Armentano, 2026 clase 7).

Matrices de Markov:
En una matriz aleatoria de Markov (columnas suman 1), siempre existe el autovalor real $\lambda = 1$ (Strang, 2019, p. I.6).

8 Matrices Ortogonales

Una matriz cuadrada $Q \in \mathbb{R}^{n\times n}$ se define como ortogonal (o más precisamente, ortonormal) si sus columnas son ortonormales (vectores unitarios perpendiculares entre sí), lo que implica que: (Armentano, 2026 clase 8; Strang, 2019, p. I.5).

$$Q^\top Q = \mathbb{I} \implies Q^\top = Q^{-1}$$

Cuando una matriz de Stiefel $Q \in \mathbb{R}^{n\times n}$ es cuadrada, sus columnas no solo son ortonormales, sino que forman una **base ortonormal** completa para $\mathbb{R}^n$. En este caso, la matriz se denomina simplemente **matriz ortogonal**.

8.0.1. La Identidad Fundamental: $Q^{-1} = Q^\top$

La propiedad más potente de una matriz ortogonal es que su transpuesta es idéntica a su inversa. [1]

Prueba: Por definición de matriz de Stiefel, $Q^\top Q = \mathbb{I}$. En el caso cuadrado, una matriz que posee una inversa a izquierda también la posee a derecha y es única. Por tanto, $Q^\top = Q^{-1}$, lo que implica que también se verifica $QQ^\top = \mathbb{I}$ (Diego Armentano, 2025 Parcial 1 Ej1a).

8.0.2. Preservación de la Geometría (Isometría)

La propiedad definitoria de las matrices ortogonales es que no alteran el producto escalar entre dos vectores cualesquiera u y v. Utilizando la propiedad de la transpuesta de un producto, $(Qu)^\top = u^\top Q^\top$, tenemos:

$$(Qu)\cdot(Qv) = (Qu)^\top(Qv) = u^\top(Q^\top Q)v = u^\top I v = u\cdot v$$

De esta invariancia se derivan dos consecuencias geométricas críticas:

1. **Preservación de la Norma (Longitud):** Dado que $\|v\| = \sqrt{v\cdot v}$, se cumple que $\|Qv\| = \|v\|$. El vector transformado mantiene su magnitud original.

[1] Esta característica simplifica drásticamente la resolución de sistemas lineales $Qx = b$. En lugar de recurrir a la eliminación gaussiana ($O(n^3)$), la solución se obtiene mediante una simple transposición $O(n^2)$: $x = Q^\top b$.

2. **Preservación de Ángulos:** El ángulo θ entre dos vectores se define por $\cos(\theta) = \frac{u \cdot v}{\|u\|\|v\|}$. Al ser el numerador y el denominador invariantes bajo Q, el ángulo permanece constante.

En definitiva, las matrices ortogonales actúan como operadores que preservan la estructura métrica del espacio. No solo mantienen las longitudes (normas), sino también los ángulos entre vectores. [2]

Por lo tanto, las matrices ortogonales $Q \in \mathbb{R}^{n \times n}$ representan una rotación (o una combinación de rotación y reflexión) porque actúan como una **isometría lineal** en el espacio euclídeo.

8.0.3. Ejemplo Canónico en $\mathbb{R}^2$

La matriz de rotación estándar para un ángulo θ en sentido antihorario es el ejemplo más claro de una matriz ortogonal:

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Podemos verificar su ortogonalidad mediante la suma de los cuadrados de sus componentes en las columnas ($\sin^2\theta + \cos^2\theta = 1$) y comprobando que su determinante es siempre 1, lo que garantiza que es una rotación pura sin reflexión.

8.0.4. El Determinante y la Orientación

Aunque todas las matrices ortogonales preservan la estructura métrica, el **determinante** de la matriz Q determina si la transformación preserva o invierte la orientación del espacio:

Transformación	Determinante	
Rotación Propia	$\det(Q) = 1$	
Rotación Impropia (Reflexión)	$\det(Q) = -1$	Invierte la orientación (quiralidad) del espacio.

[2] Las transformaciones basadas en matrices ortogonales (como las rotaciones de Householder o Givens) son preferidas en aplicaciones donde la estabilidad numérica es crítica porque **nunca amplifican el error ni producen overflow**. Dado que $\|Qx\| = \|x\|$, el tamaño de los datos se mantiene constante a lo largo de múltiples factorizaciones.

8.0.5. Ejemplos Clave en Computación

Matrices de Permutación (P):
Reordenan las entradas de un vector. Sus columnas son las de la identidad en distinto orden. Siempre verifican $P^\top = P^{-1}$.

Matrices de Rotación:
En $\mathbb{R}^2$, la matriz $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ rota vectores preservando el origen. Sus autovalores son complejos: $\lambda = \cos\theta \pm i\sin\theta$.

Matrices de Refleccion:
En $\mathbb{R}^2$, la matriz $Q = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$ refleja los vectores respecto de la recta que pasa por el origen y tiene un angulo θ con el eje horizontal.

Matriz de Fourier (F):
Utilizada en el procesamiento de señales, sus columnas son autovectores ortonormales que permiten transitar entre el dominio del tiempo y la frecuencia.

8.1. Espectro

Los valores propios (λ) de una matriz ortogonal siguen reglas geométricas estrictas derivadas de su capacidad para preservar longitudes.

8.1.0.1. Módulo Unitario ($|\lambda| = 1$)

Debido a que las matrices ortogonales preservan la norma de los vectores ($\|Qx\| = \|x\|$) (Strang, 2018 Lec. 3 '62). Si aplicamos esto a un autovector:

$$\|Qx\| = \|\lambda x\| = |\lambda|\|x\| \implies |\lambda| = 1$$

Por lo tanto, todos los autovalores de una matriz ortogonal deben tener magnitud 1; es decir, residen en el círculo unidad del plano complejo (Strang, 2019, p. I.5).

8.1.0.2. Valores Reales Restringidos

Los únicos valores propios reales que puede tener una matriz ortogonal son 1 o -1. Cualquier otro valor real violaría la condición de módulo uno.

8.1.0.3. Simetría en el Plano Complejo

Como Q tiene entradas reales, sus autovalores complejos siempre aparecen en pares λ y $\bar{\lambda}$ (Teorema Fundamental del Álgebra). Esto asegura que el polinomio característico mantenga coeficientes reales. Esto puede verse algebraicamente: Si λ es un valor propio complejo de Q con vector propio v, entonces: $Q\overline{v} = \overline{Qv} = \overline{\lambda v} = \overline{\lambda}\,\overline{v}$. Por lo tanto, $\overline{\lambda}$ también es valor propio con vector propio $\overline{v}$.

8.1.0.4. Determinante e Invarianza

El determinante de Q, que es el producto de sus valores propios, siempre es 1 (rotaciónes puras) o -1 (reflexión). A partir de la identidad $Q^\top Q = I$, aplicamos la propiedad del determinante de un producto: $\det(Q^\top)\det(Q) = \det(I)$. Sabiendo que $\det(Q^\top) = \det(Q)$ y que $\det(I) = 1$, obtenemos la relación $(\det(Q))^2 = 1$

8.1.0.5. Espectro de $Q^\top = Q^{-1}$

Como se demostró en Sección 7.4, Q y $Q^\top$ siempre comparten los mismos autovalores. Sin embargo, en el caso de las matrices ortogonales, existe una relación adicional: $Q^\top = Q^{-1}$. Esto implica que: **los autovalores de $Q^\top$ son también los autovalores del inverso de Q**.

Si x es un autovector de Q con autovalor λ, entonces (Strang, 2018 Lec. 4 '115):

$$Qx = \lambda x \implies x = Q^{-1}(\lambda x) \implies Q^{-1}x = \frac{1}{\lambda}x$$

Por lo tanto: **si λ es un autovalor de Q, su recíproco $1/\lambda$ es un autovalor de Q^{-1} (y por ende de $Q^\top$)**.

> **i Nota**
>
> Para cualquier número complejo λ con $|\lambda| = 1$, se cumple que su recíproco es igual a su conjugado:
>
> $$\frac{1}{\lambda} = \bar{\lambda}$$
>
> Los autovalores de $Q^\top$ son los recíprocos de los de Q, pero como tienen magnitud 1, los autovalores de la traspuesta son simplemente los **complejos conjugados** de los autovalores originales (Strang, 2019, p. I.5).

8.2. Vectores Propios

Al igual que las matrices simétricas, las matrices ortogonales (y en general las matrices normales) poseen vectores propios que son ortogonales entre sí. En el caso de valores propios complejos, esta ortogonalidad se verifica utilizando el producto interno complejo $x_1^H x_2 = 0$.

8.3. Ejemplo: Matriz de Rotación en $\mathbb{R}^2$

Consideramos la matriz de rotación estándar Q (Hernández, 2026, p. P3, ejercicio 7), que gira cualquier vector del plano un ángulo θ

$$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Una matriz de rotación pura (con $\theta \neq 0, \pi$) en $\mathbb{R}^2$ típicamente no tiene autovectores reales porque gira todos los vectores un ángulo θ. Sin embargo, posee autovalores complejos de la forma $\lambda = e^{\pm i\theta} = \cos\theta \pm i\sin\theta$, que cumplen con tener magnitud 1. $|e^{\pm i\theta}| = 1$. Solo en los casos donde $\theta = 0$ ($\lambda = 1$) o $\theta = \pi$ ($\lambda = -1$) los vectores propios vuelven al espacio real.

Los autovectores asociados son vectores complejos en $\mathbb{C}^2$:

- Para $\lambda_1 = e^{i\theta}$ el autovector es $x_1 = (1, -i)$.
- Para $\lambda_2 = e^{-i\theta}$ el autovector es $x_2 = (1, i)$.

Podemos validar estos resultados mediante la traza y el determinante:

- **Traza:** $\lambda_1 + \lambda_2 = (\cos\theta + i\sin\theta) + (\cos\theta - i\sin\theta) = 2\cos\theta$, que coincide exactamente con la suma de la diagonal de Q.
- **Determinante:** $\lambda_1 \cdot \lambda_2 = e^{i\theta} \cdot e^{-i\theta} = e^0 = 1$, consistente con una rotación que preserva la orientación.

> **i** Ortogonalidad Compleja
>
> Aunque estos vectores viven en $\mathbb{C}^2$, siguen siendo ortogonales entre sí bajo el producto interno complejo $x_1^H x_2 = 0$.

8.4. Teorema Espectral para Matrices Simétricas

Sea $S \in \mathcal{M}^{n\times n} : S = S^\top$, se cumple:

1. **Autovalores Reales:** Sus autovalores son todos reales ($\lambda_i \in \mathbb{R}$).

2. **Ortogonalidad de Autovectores:** Los autovectores correspondientes a autovalores distintos son ortogonales entre si.
3. **Diagonalización Ortogonal:** Existe siempre una base ortonormal de autovectores, lo que permite una diagonalización mediante una **matriz ortogonal** Q: [3]

$$S = Q\Lambda Q^\top = \sum_{i=1}^{n} \lambda_i q_i q_i^\top$$

8.4.1. Prueba: Ortogonalidad de Autovectores

Sean q_1, q_2 autovectores de S con $\lambda_1 \neq \lambda_2$.

$$Sq_1 = \lambda_1 q_1 \implies (Sq_1)^\top = \lambda_1 q_1^\top \implies q_1^\top S = \lambda_1 q_1^\top$$

Multiplicando por q_2 a la derecha:

$$q_1^\top S q_2 = \lambda_1 q_1^\top q_2$$

Sustituyendo $Sq_2 = \lambda_2 q_2$:

$$\lambda_2 q_1^\top q_2 = \lambda_1 q_1^\top q_2 \implies (\lambda_2 - \lambda_1) q_1^\top q_2 = 0$$

Dado que $\lambda_1 \neq \lambda_2$, necesariamente $q_1^\top q_2 = 0$. Por lo tanto, los autovectores son ortogonales.

8.4.2. Prueba: Descomposición Espectral como Suma de Rango 1

La identidad $S = Q\Lambda Q^\top$ puede expandirse utilizando la multiplicación columna-fila como una suma de matrices de **rango 1**:

$$S = \lambda_1 q_1 q_1^\top + \lambda_2 q_2 q_2^\top + \cdots + \lambda_n q_n q_n^\top$$

Cada término $\lambda_i q_i q_i^\top$ representa la proyección del espacio sobre la dirección del autovector q_i, escalada por λ_i. [4]

[3] Esta expansión como **suma de matrices de rango 1** es fundamental para entender la descomposición de la varianza en estadística.

[4] En aplicaciones de ciencia de datos, esta expansión permite identificar las componentes que capturan la mayor varianza del sistema, descartando aquellas con autovalores despreciables.

8.5. Aplicaciónes

8.5.1. Transformaciones y Cálculo Funcional

La factorización $A = X\Lambda X^{-1}$ simplifica el cálculo de funciones de matrices:

1. **Potencias:** $A^k = X\Lambda^k X^{-1}$. Si $|\lambda_i| < 1$, entonces $A^k \to 0$ cuando $k \to \infty$.
2. **Inversa:** $A^{-1} = X\Lambda^{-1}X^{-1}$ (si $\lambda_i \neq 0$).
3. **Desplazamiento (Shift):** Los autovectores de $A + sI$ son los mismos que los de A, pero sus autovalores son $\lambda_i + s$.

8.5.2. Reducción de Dimensionalidad Mediante el Método de la Potencia

Desde una perspectiva numérica, si deseamos encontrar la dirección del autovector dominante (el asociado al $|\lambda|$ más grande), podemos iterar un vector aleatorio v_0:

$$v_{k+1} = \frac{Av_k}{\|Av_k\|}$$

A medida que $k \to \infty$, el vector v_k se alinea con el autovector dominante.

9 Matrices Definidas y Semidefinidas Positivas

Las matrices definidas positivas representan una clase especial de matrices simétricas que desempeñan un papel central en la optimización, la estadística y el análisis de estabilidad de sistemas. Su estudio permite pasar de una visión puramente lineal a una comprensión de superficies cuadráticas y "energía" del sistema.

9.1. Definición

(Armentano, 2026 Clase 9) Sea $S \in \mathbb{R}^{n \times n}$ una matriz **simétrica** ($S = S^\top$). Decimos que S es una **matriz definida positiva** si para todo vector no nulo $x \in \mathbb{R}^n$, el número escalar resultante de la forma cuadrática es estrictamente mayor que cero:

$$x^\top S x > 0 \quad \forall x \neq 0$$

Geométricamente, la expresión $f(x) = x^\top S x$ define una superficie en $\mathbb{R}^n$. Si la matriz es DP, la gráfica de $f(x)$ se asemeja a un paraboloide que abre hacia arriba (bowl), con un mínimo global único en el origen $x = 0$.

Dependiendo del comportamiento de la operación $x^\top S x$, clasificamos la matriz simétrica S en:

- **Definida Positiva:** Si $x^\top S x > 0$ para todo $x \neq 0$.
- **Semidefinida Positiva:** Si $x^\top S x \geq 0$ para todo x. Esto permite que existan vectores no nulos (normalmente en el núcleo de la matriz) para los cuales la forma cuadrática sea cero.
- **Indefinida:** Si existen vectores x tales que $x^\top S x > 0$ y otros vectores y tales que $y^\top S y < 0$.

Nomenclatura:

- Utilizamos tanto las abreviaciones en inglés: **PD** = Positive Definite, **PSD** = Positive Semi-Definite, como en español: **MDP** = Matriz Definida Positiva, **MSDP** = Matriz Semi-Definida Positiva.
- Al número resultante de la operación $x^\top S x$ se le denomina **Energía** o **Forma Cuadrática**.

9.2. Criterios de verificación

Para verificar si una matriz simétrica S es PD, existen diversos criterios equivalentes que vinculan la forma cuadrática con el espectro de la matriz y sus factorizaciones. (Strang, 2018 Lec 5)

9.2.1. 1: Autovalores Positivos

Una matriz simétrica S es definida positiva si y solo si todos sus autovalores son estrictamente positivos:

$$\lambda_i > 0 \quad \text{para todo } i = 1, \dots, n$$

El teorema espectral nos dice que existe una base de vectores propios que diagonaliza a S como $S = Q\Lambda Q^\top$. La forma cuadrática se puede escribir entonces como $x^\top(Q\Lambda Q^\top)x$. Definiendo $y = Q^\top x$ la expresion se simplifica a $S = y^\top \Lambda y = \sum \lambda_i y_i^2$, donde es facil ver que si todos los $\lambda_i > 0$, la suma es necesariamente positiva para cualquier $y \neq 0$.

En el caso de una matriz S **Semidefinida Positiva** se permite que algunos autovalores sean cero ($\lambda_i \geq 0$), lo que implica que la matriz puede ser singular.

9.2.2. 2: Energía Positiva

Strang usa el concepto de Energía positiva, que no es más que la definicion que usamos aquí.

9.2.3. 3: Factorización $S = A^\top A$

Si S es PD, existe una matriz invertible A tal que $S = A^\top A$ [1].

9.2.4. 4: Determinante y Menores Principales Positivos

Una matriz S es PD si todos sus **menores principales líderes** son positivos. Esto incluye:

- El determinante de la matriz completa: $\det(S) = \prod \lambda_i > 0$.
- Todos los determinantes de las submatrices superiores izquierdas de tamaño $k \times k$.

[1] En computación numérica, esto se traduce en la **Factorización de Cholesky**: $S = LL^\top$, donde L es una matriz triangular inferior con diagonal positiva.

9.2.5. 5: Pivotes de eliminacion Gaussiana Positivos

Toda matriz simétrica definida positiva permite una eliminación gaussiana sin necesidad de intercambio de filas (pivoteo), lo que conduce a la **descomposición simétrica** $A = LDL^\top$. En esta estructura:

- L Es una matriz triangular inferior unitaria (con 1s en la diagonal principal). Representa las operaciones elementales de fila realizadas durante la eliminación.
- D: Es una matriz diagonal que contiene los **pivotes** resultantes del proceso de eliminación: $D = \text{diag}(d_1, d_2, \dots, d_n)$.
- $L^\top$: Es la transpuesta de L, una matriz triangular superior.

Para demostrar por qué los pivotes deben ser positivos, analizamos la "energía" o **forma cuadrática** asociada a la matriz. Comenzamos por multiplicar por un vector no nulo $x \in \mathbb{R}^n$ en ambos lados:

$$x^\top Ax = x^\top(LDL^\top)x$$

Por la propiedad asociativa del producto matricial, podemos agrupar los términos de la siguiente manera:

$$x^\top Ax = (L^\top x)^\top D(L^\top x)$$

Si definimos un nuevo vector columna y como $y = L^\top x$, la expresión se simplifica a una suma ponderada de cuadrados:

$$x^\top Ax = y^\top Dy = \sum_{i=1}^{n} d_i y_i^2 = d_1 y_1^2 + d_2 y_2^2 + \dots + d_n y_n^2$$

Dado que L (y por ende $L^\top$) es una matriz triangular con 1s en la diagonal, su determinante es 1 y es, por tanto, invertible. Esto garantiza que para cualquier $y \neq 0$, siempre existe un único $x \neq 0$ tal que $L^\top x = y$. Para que la suma $\sum_{i=1}^{n} d_i y_i^2$ sea estrictamente positiva para cualquier valor de y_i (donde al menos un $y_i \neq 0$), **todos los coeficientes d_i deben ser estrictamente mayores a cero**.

Si existiera un pivote $d_k \leq 0$, podríamos elegir un vector y con componentes nulas excepto en la posición k, lo que resultaría en una forma cuadrática $x^\top Ax \leq 0$, invalidando la propiedad de ser definida positiva.

9.3. Ejemplos y Aplicaciones Fundamentales

Para identificar la positividad en la práctica, analizamos estructuras matriciales recurrentes en el álgebra lineal numérica y la ciencia de datos, vinculándolas con la estructura de los subespacios y la teoría espectral.

9.3.1. Relación con la Matriz Identidad

Cualquier matriz definida positiva S puede verse como una deformación de la matriz identidad $\mathbb{I}$. Mientras que $x^\top \mathbb{I} x = \|x\|^2$ define una esfera perfecta (todas las direcciones pesan lo mismo), $x^\top S x$ define un elipsoide donde los ejes están alineados con los autovectores de S y sus longitudes dependen de $1/\sqrt{\lambda_i}$.

Un resultado útil es que si S es PD, entonces existe un escalar $\alpha > 0$ tal que: $S - \alpha\mathbb{I}$ es semidefinida positiva. Esto indica que S "crece" al menos tan rápido como una versión escalada de la identidad.

9.3.2. Matrices de Gram ($A^\top A$)

Para cualquier matriz $A \in \mathbb{R}^{m\times n}$, el producto $S = A^\top A$ siempre es PSD. $A^\top A$ es estrictamente PD si y solo si las columnas de A son linealmente independientes (rango n). En ese caso, $Ax = 0$ implica necesariamente $x = 0$.

Prueba: La forma cuadrática se reduce a la norma al cuadrado del vector transformado:

$$x^\top (A^\top A)x = (Ax)^\top (Ax) = \|Ax\|^2$$

Dado que una norma siempre es ≥ 0, el resultado es siempre no negativo.

9.3.3. La Matriz de Unos (J)

La matriz J, donde cada entrada $J_{ij} = 1$, es un caso crítico de matriz de **rango 1**. Puede expresarse como el producto exterior $J = \mathbf{1}\mathbf{1}^\top$, donde $\mathbf{1}$ es el vector columna de unos. Sus autovalores son $\lambda = \{n, 0, 0, \ldots, 0\}$. Al tener autovalores no negativos, es **semidefinida positiva**. Es singular para cualquier $n > 1$ porque su núcleo tiene dimensión $n - 1$.

9.3.4. Matrices Simétricas, Espectro Real, Matrices de Proyección Ortogonal

Toda matriz definida positiva debe ser, por definición, simétrica. La simetría garantiza que todos los autovalores sean reales, lo que permite ordenarlos y validar si el autovalor mínimo $\lambda_{min} > 0$. Un ejemplo común son las matrices de proyección ortogonal $P = QQ^\top$, las cuales son PSD pero no PD (tienen autovalores 1 y 0).

9.3.5. Matrices de Markov Simétricas

Una matriz de Markov M cumple que sus columnas suman 1. Aunque no todas las matrices de Markov son PSD, si además M es simétrica, sus autovalores son reales y el autovalor dominante es $\lambda_1 = 1$.

9.3.6. Desplazamientos (Shifts)

Una matriz simétrica arbitraria S puede convertirse en definida positiva mediante un “shift” o desplazamiento espectral utilizando la matriz identidad $\mathbb{I}$ (Armentano (2026) clase 8). Si λ_{min} es el autovalor más pequeño de S, entonces para cualquier $s > |\lambda_{min}|$ se cumple que $S + s\mathbb{I}$ es estrictamente Definida Positiva.

Referencias

Armentano, D. (2026). *Notas de clase del curso Geometría y Álgebra Lineal para Aprendizaje Estadístico*. Facultad de Ciencias Económicas y Administracion, Universidad de la República. https://github.com/diegoax/ALNAE-2026/tree/main/Notebooks

Diego Armentano, M. H. (2025). *Pruebas parciales de Álgebra Lineal para Aprendizaje Estadístico*. Facultad de Ciencias Económicas y Administracion, Universidad de la República. https://github.com/diegoax/ALNAE-2026/tree/main

Hernández, M. (2026). *Ejercicios practicos del curso Geometría y Álgebra Lineal para Aprendizaje Estadístico*. Facultad de Ciencias Económicas y Administracion, Universidad de la República. https://github.com/diegoax/ALNAE-2026/tree/main/practicos

Strang, G. (2018). *Matrix Methods in Data Analysis, Signal Processing, and Machine Learning*. Massachusetts Institute of Technology: MIT OpenCourseWare. https://ocw.mit.edu/courses/18-065-matrix-methods-in-data-analysis-signal-processing-and-machine-learning-spring-2018/

Strang, G. (2019). *Linear Algebra and Learning from Data*. Wellesley-Cambridge Press. https://math.mit.edu/~gs/learningfromdata/

A Eliminación Gaussiana y la Factorización LU

La eliminación gaussiana es el algoritmo fundamental para la resolución de sistemas lineales $Ax = b$. Desde la perspectiva del álgebra lineal numérica, este proceso se interpreta como la factorización de la matriz original en dos componentes triangulares, permitiendo una comprensión profunda de la propagación de la información dentro del sistema.

Un paradigma moderno y sumamente útil consiste en visualizar cada etapa de la eliminación gaussiana como la resta de una **matriz de rango 1** a la matriz de trabajo actual.

Sea $A \in \mathbb{R}^{n \times n}$. La matriz puede expresarse como la suma de n matrices de rango 1, donde cada término representa el aporte de un paso de eliminación:

$$A = \sum_{i=1}^{n} \ell_i u_i^T = \ell_1 u_1^T + \ell_2 u_2^T + \cdots + \ell_n u_n^T$$

donde:

- ℓ_i es un vector columna que contiene los multiplicadores aplicados debajo del i-ésimo pivote.
- u_i^T es el vector fila correspondiente a la fila pivote original en esa etapa.

Este enfoque revela que la eliminación reduce sistemáticamente la complejidad del sistema, "extrayendo" una dirección independiente en cada paso hasta agotar el rango de la matriz.

A.1. Factorización $A = LU$

Si el proceso de eliminación se completa sin necesidad de intercambios de filas, la matriz A admite una factorización única de la forma:

$$A = LU$$

Matriz L (Lower):
Es una matriz **triangular inferior** con unidades en la diagonal principal ($L_{ii} = 1$). Sus entradas por debajo de la diagonal son precisamente los multiplicadores utilizados para eliminar las variables.

Matriz U (Upper):
Es una matriz **triangular superior** que contiene los coeficientes resultantes tras completar la eliminación. Sus entradas diagonales son los **pivotes**.

La matriz L se construye agrupando los vectores columna ℓ_i, y la matriz U agrupando los vectores fila u_i^T:

$$L = [\ell_1, \ell_2, \dots, \ell_n], \quad U = \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_n^T \end{bmatrix}$$

Es vital notar que para que L tenga unos en la diagonal, los vectores ℓ_i deben estar normalizados respecto al valor del pivote en ese paso.

A.2. Factorización $PA = LU$

En la práctica numérica, la eliminación puede fallar si se encuentra un cero en la posición del pivote, o ser inestable si el pivote es muy pequeño. Para solventar esto, se recurre al **pivoteo parcial**.

Teorema de Existencia General:

Para toda matriz A invertible, existe una **matriz de permutación** P tal que la matriz reordenada admite una factorización LU:

$$PA = LU$$

La matriz P simplemente registra los intercambios de filas necesarios para asegurar que los pivotes sean los máximos posibles en valor absoluto, minimizando el error de redondeo.

A.3. Complejidad Computacional

La multiplicación de matrices y su factorización involucran un conteo de operaciones fundamental para el análisis de algoritmos. Para matrices de tamaño $n \times n$:

1. **Eliminación:** El costo es de aproximadamente $\frac{1}{3}n^3$ multiplicaciones y sumas.
2. **Resolución:** Una vez obtenida la forma LU, resolver $Ly = b$ (sustitución hacia adelante) y $Ux = y$ (sustitución hacia atrás) requiere solo n^2 operaciones.

Esta asimetría en el costo subraya la importancia de la factorización: una vez que A es factorizada, resolver para múltiples vectores b es extremadamente eficiente.

B Descomposición $S = Q\Lambda Q^T$

(Strang, 2018 Lec 2 '4)

La descomposición $Q\Lambda Q^T$ es una variante específica de la **diagonalización ortogonal** aplicada exclusivamente a **matrices simétricas**. En el álgebra lineal numérica, esta forma es fundamental porque garantiza propiedades computacionales y teóricas que no están presentes en matrices generales.

Para una matriz cuadrada $A \in \mathbb{R}^{n\times n}$ que es simétrica (es decir, $A = A^T$), el **Teorema Espectral** establece que A es ortogonalmente diagonalizable. La expresión se define formalmente como:

$$A = Q\Lambda Q^T$$

Donde:

- Q: Es una **matriz ortogonal** cuyas columnas son los autovectores (vectores propios) de A. Al ser ortogonal, cumple la propiedad $Q^T = Q^{-1}$, lo que implica que $QQ^T = I$.
- Λ: Es una **matriz diagonal** que contiene los autovalores (valores propios) $\lambda_1, \lambda_2, \dots, \lambda_n$ correspondientes a cada autovector en Q.

B.1. Componentes de la Descomposición

B.1.0.1. La Matriz de Autovalores (L)

En la matriz diagonal Λ, todos los elementos fuera de la diagonal principal son nulos. La importancia de Λ radica en que representa el factor de escala de la transformación lineal a lo largo de sus ejes principales.

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

B.1.0.2. La Matriz Ortogonal (Q)

Las columnas de Q forman una **base ortonormal** del espacio euclídeo. Esto implica que cada columna tiene norma unitaria ($\|q_i\| = 1$) y es ortogonal a las demás ($q_i \cdot q_j = 0$ para $i \neq j$). Geométricamente, Q opera como una **isometría** (rotación o reflexión).

B.2. Interpretación Geométrica

Si visualizamos la matriz A como una transformación lineal, la descomposición $Q\Lambda Q^T$ desglosa la operación en tres etapas cinemáticas:

1. Q^T: Realiza un cambio de base del vector de entrada al sistema de coordenadas definido por los autovectores (rota el espacio).
2. Λ: Aplica un escalamiento (dilatación o contracción) en cada una de estas nuevas direcciones según los valores λ_i.
3. Q: Devuelve el vector al sistema de coordenadas original mediante la rotación inversa.

B.3. Importancia y Aplicaciones Técnicas

- **Cálculo de Potencias de Matrices:** Elevar una matriz simétrica a una potencia k se simplifica operacionalmente: $A^k = Q\Lambda^k Q^T$. Solo se requiere elevar los elementos escalares de la diagonal de Λ.
- **Análisis de Componentes Principales (PCA):** En estadística multivariante, la matriz de covarianza es simétrica y semidefinida positiva. Su descomposición permite identificar las direcciones de máxima varianza.
- **Estabilidad Numérica:** El uso de matrices ortogonales es preferible en computación científica debido a que poseen un **número de condición** óptimo, evitando la amplificación de errores de redondeo (Trefethen & Bau, 1997).

C Descomposición en Valores Singulares (SVD)

(Strang, 2018 Lec. 6; Strang, 2019, p. I.7)

A diferencia de la diagonalización, que requiere que la matriz sea cuadrada y posea suficientes autovectores, y la descomposición espectral $S = Q\Lambda Q^T$ que se limita a matrices simétricas, la SVD es una generalizacion universal: existe para cualquier matriz $A \in \mathbb{R}^{m\times n}$, ya sea cuadrada o rectangular.

C.1. 10.1 Motivación: Limitaciones de los Autovectores

La noción de vectores propios ($Ax = \lambda x$) presenta algunos inconvenientes:

1. **Matrices Rectangulares:** Si A es $m \times n$, el producto Ax reside en $\mathbb{R}^m$, mientras que x está en $\mathbb{R}^n$, haciendo imposible la igualdad $Ax = \lambda x$.
2. **Matrices Cuadradas Generales:** Incluso si A es cuadrada, los autovalores pueden ser complejos o los autovectores pueden no ser ortogonales.

La SVD supera esto utilizando dos conjuntos distintos de vectores ortonormales: los **vectores singulares a izquierda (U)** y a **derecha (V)**.

C.2. La Ecuación Fundamental

Toda matriz A puede factorizarse en el producto de tres matrices:

$$A = U\Sigma V^T$$

Donde:

- $U \in \mathbb{R}^{m\times m}$: Es una matriz ortogonal cuyas columnas son los **vectores singulares a izquierda** y representan una base ortonormal para el codominio
- $\Sigma \in \mathbb{R}^{m\times n}$: Es una matriz diagonal (aunque rectangular) que contiene los **valores singulares** $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$
- $V \in \mathbb{R}^{n\times n}$: Es una matriz ortogonal cuyas columnas son los **vectores singulares a derecha** y representan una base ortonormal para el dominio.

C.3. Intuición Geométrica: Rotación y Estiramiento

La acción de multiplicar por una matriz A puede descomponerse en tres pasos geométricos elementales (Strang, 2018 Lec. 3 '6):

1. **Rotación en el dominio (V^T):** Cambia el vector de entrada a la base ortonormal de vectores singulares sin cambiar su longitud].
2. **Escalamiento (Σ):** Estira o contrae el vector a lo largo de los ejes principales. Aquí, la esfera unitaria se transforma en un elipsoide cuyos semiejes tienen longitudes σ_i.
3. **Rotación en el codominio (U):** Orienta el elipsoide resultante en el espacio final.

C.4. De las Ecuaciones Vectoriales a la Factorización Matricial

Para construir la SVD, buscamos una relación análoga a la de los autovectores ($Ax = \lambda x$), pero adaptada a matrices rectangulares donde el dominio y el codominio tienen dimensiones distintas.

El punto de partida es encontrar un conjunto de vectores ortonormales en el dominio $\{v_1, v_2, \dots, v_r\}$ que, al ser transformados por A, resulten en vectores ortonormales en el codominio $\{u_1, u_2, \dots, u_r\}$, escalados por un factor σ_i:

$$Av_i = \sigma_i u_i \quad , i = 1, \dots, r$$

Donde r es el rango de la matriz. Estas ecuaciones representan la acción de la matriz sobre sus direcciones principales (Strang, 2018 - Lecture 6 '20).

Podemos agrupar estas r ecuaciones individuales colocando los vectores v_i como columnas de una matriz V y los vectores u_i como columnas de una matriz U (Strang, 2018 - Lecture 6 '30):

$$A \underbrace{\begin{bmatrix} | & | & & | \\ v_1 & v_2 & \dots & v_r \\ | & | & & | \end{bmatrix}}_{V} = \underbrace{\begin{bmatrix} | & | & & | \\ u_1 & u_2 & \dots & u_r \\ | & | & & | \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}}_{\Sigma}$$

Esta disposición compacta las relaciones vectoriales en la ecuación:

$$AV = U\Sigma$$

Dado que los vectores v_i son elegidos para ser ortonormales, la matriz V cumple la propiedad de las matrices de Stiefel: $V^T V = \mathbb{I}$ (Strang, 2018 Lec. 3 '02; Armentano, 2026 clase 5).

Al multiplicar por V^T a la derecha en ambos lados de la ecuación anterior, obtenemos la representación final de la SVD $A = U\Sigma V^T$

$$A = \underbrace{\begin{bmatrix} | & | & & | \\ u_1 & u_2 & \dots & u_r \\ | & | & & | \end{bmatrix}}_{U \in \mathcal{M}_{m\times r}} \underbrace{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}}_{\Sigma \in \mathcal{M}_{r\times r}} \underbrace{\begin{bmatrix} - & v_1^T & - \\ - & v_2^T & - \\ & \vdots & \\ - & v_r^T & - \end{bmatrix}}_{V^T \in \mathcal{M}_{r\times n}}$$

En esta estructura, las columnas de U representan la base ortonormal de la imagen, mientras que las **filas de V^T** (que son los vectores columna v_i transpuestos) representan la base ortonormal del espacio fila.

Entonces, la conexión con la estructura espectral de las matrices simétricas permite identificar la procedencia de estos vectores:

- **Las columnas de V:** Provienen de la diagonalización de $A^T A$, siendo sus vectores propios.
- **Las columnas de U:** Provienen de la diagonalización de AA^T, siendo sus vectores propios.

C.5. Relación con la Estructura Espectral

La construcción de la SVD se apoya en las matrices simétricas $A^T A$ y AA^T, las cuales siempre son semidefinidas positivas

- **$A^T A$:** Es una matriz cuadrada ($n \times n$), simétrica y **definida positiva**. Sus autovectores forman las columnas de V y sus autovalores son σ_i^2.
- **AA^T:** Es una matriz cuadrada ($m \times m$) que posee los mismos autovalores no nulos que $A^T A$. Sus autovectores forman las columnas de U.

1. **Vectores singulares a derecha (V)**: Son los autovectores de la matriz simétrica $A^T A$. Se cumple que $(A^T A)v_i = \sigma_i^2 v_i$
2. **Vectores singulares a izquierda (U)**: Son los autovectores de la matriz simétrica AA^T. Se cumple que $(AA^T)u_i = \sigma_i^2 u_i$
3. **Valores singulares (σ_i)**: Son las raíces cuadradas de los autovalores no nulos de $A^T A$ (o AA^T). Representan el factor de estiramiento en cada dirección invariante.

Algebraicamente, esto se verifica expandiendo el producto:

$$A^T A = (V\Sigma^T U^T)(U\Sigma V^T) = V(\Sigma^T \Sigma)V^T$$

Dado que $U^T U = \mathbb{I}$, se observa que V diagonaliza a $A^T A$, y en esta descomposición se ve que las columnas de V son los vectores propios de $A^T A$

C.6. Construcción y Prueba de Ortogonalidad

Buscamos vectores que cumplan la relación $Av_i = \sigma_i u_i$.

1. Se eligen los v_i como los autovectores de $A^T A$.
2. Se calculan los σ_i como la raíz cuadrada de los autovalores de $A^T A$.
3. Se definen los vectores de salida como $u_i = \frac{Av_i}{\sigma_i}$.

Para demostrar que los u_i generados son ortogonales entre sí ($u_1^T u_2 = 0$), utilizamos la propiedad de los autovectores de $A^T A$:

$$u_1^T u_2 = \left(\frac{Av_1}{\sigma_1}\right)^T \left(\frac{Av_2}{\sigma_2}\right) = \frac{v_1^T (A^T A v_2)}{\sigma_1 \sigma_2} = \frac{v_1^T (\sigma_2^2 v_2)}{\sigma_1 \sigma_2} = \frac{\sigma_2}{\sigma_1}(v_1^T v_2)$$

Como los v_i son ortonormales, $v_1^T v_2 = 0$, lo que garantiza la ortogonalidad de los u_i.

C.7. Observaciones y Propiedades

C.7.0.1. Conteo de Parámetros

La SVD describe cualquier transformación lineal como una secuencia de **Rotación (V^T)** $\rightarrow$ **Estiramiento (Σ)** $\rightarrow$ **Rotación (U)**.

El número de grados de libertad en la matriz original debe coincidir con los de su forma SVD:

- **Caso 2x2 (4 parámetros):** 2 valores singulares en Σ, 1 ángulo de rotación θ para V y 1 ángulo para U.
- **Caso 3x3 (9 parámetros):** 3 valores singulares en Σ, 3 parámetros de rotación en 3D (los tres angulos posibles, conocidos en ingeniería como *roll, pitch* y *yaw*) para V, y otros 3 para U.

C.7.0.2. Igualdad de U y V

Los vectores singulares a izquierda y derecha coinciden ($U = V$) únicamente cuando la matriz A es cuadrada, simétrica y definida positiva.

C.7.0.3. Autovalores Repetidos

Si existe un autovalor doble, los vectores singulares no son únicos, sino que generan un **plano singular** de direcciones posibles

C.7.0.4. Determinantes

El determinante de una matriz ortogonal es siempre 1 (o -1 si incluye reflexión) (Strang, 2018 Lec. 3 '13). Para una matriz cuadrada, el producto de los valores singulares $\prod \sigma_i$ es igual al valor absoluto del determinante.

C.8. Aproximación de Bajo Rango (Teorema de Eckart-Young)

En ciencia de datos, la SVD permite comprimir información eliminando los componentes menos significativos (ruido). Según el Teorema de Eckart-Young, la mejor aproximación de rango k para una matriz A se obtiene truncando la SVD a sus primeros k términos:

$$A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^T$$

Esta matriz A_k es la más cercana a A en términos de la norma de Frobenius y la norma inducida ℓ_2 (Strang, 2018 Lec. 7 '16; Strang, 2019, p. I.7).

C.9. La Pseudo-inversa (A^+)

La SVD proporciona una forma robusta de definir la inversa para matrices no invertibles o rectangulares (Strang, 2018 Lec. 9 '06; Strang, 2019, p. I.7):

$$A^+ = V\Sigma^+ U^T$$

Donde Σ^+ se obtiene reemplazando cada $\sigma_i > 0$ por $1/\sigma_i$ y transponiendo la matriz. Si A tiene columnas independientes, $A^+ = (A^T A)^{-1} A^T$, lo que conecta directamente con la solución de mínimos cuadrados.