

Series de Tiempo con Bosques Aleatorios (Random Forest)

Wilson Eduardo Jeréz-Hernández

2 Diciembre de 2023

Series de Tiempo

Series de Tiempo

- Una serie de tiempo es un conjunto de datos que se recopila o registra en intervalos regulares a lo largo del tiempo.
- Estas series pueden mostrar patrones complejos, tendencias a largo plazo y variaciones estacionales que las hacen adecuadas para una amplia gama de aplicaciones, desde la predicción financiera hasta el monitoreo de la salud.

Series de Tiempo

Series de Tiempo



Figure 1: <https://www.pricing.cl/conocimiento/series-de-tiempo/>

Características de una Serie de Tiempo

- ① Movimientos a Largo Plazo o Tendencia.
- ② Movimientos Cíclicos: Fluctuaciones que ocurren en patrones recurrentes, pero no necesariamente a intervalos fijos.
- ③ Movimientos Estacionales: Intervalos regulares a lo largo del tiempo.
- ④ Movimientos Irregulares o Aleatorios.

Series de Tiempo

Características de una Serie de Tiempo

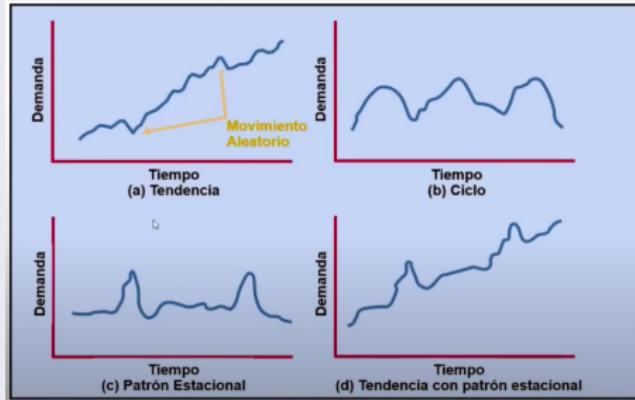


Figure 2: <https://www.youtube.com/watch?app=desktop&v=aUwXWGa8jK0>

Tipos de Análisis en Series de Tiempo

Existen diferentes enfoques para analizar series de tiempo, dependiendo de los objetivos de la investigación. Los tres tipos principales de análisis son:

- ① Análisis Descriptivo.
- ② Análisis Exploratorio/Inferencial.
- ③ Análisis Predictivo.

Análisis Descriptivo.

- ① Este tipo de análisis tiene como objetivo comprender y describir las características fundamentales de una serie de tiempo.
- ② Proporciona una visión general de la tendencia, ciclos, estacionalidad y movimientos irregulares presentes en los datos.
- ③ Ejemplos incluyen identificar patrones de comportamiento, calcular estadísticas descriptivas como promedios y desviaciones estándar, y visualizar la serie a través de gráficos.

Análisis Exploratorio/Inferencial.

- ① El análisis exploratorio profundiza en la identificación y comprensión de los factores subyacentes que influyen en el comportamiento de la serie de tiempo.
- ② Busca descubrir relaciones causa-efecto o correlaciones entre variables que puedan explicar patrones observados.
- ③ Ejemplos incluyen el uso de técnicas estadísticas avanzadas, como el análisis de regresión, para modelar y comprender las relaciones entre variables.

Análisis Predictivo.

- ① El análisis predictivo utiliza datos históricos para hacer proyecciones sobre el comportamiento futuro de la serie de tiempo.
- ② Se basa en modelos y algoritmos de predicción que pueden prever valores futuros, tendencias y patrones.
- ③ Ejemplos incluyen el uso de modelos de series de tiempo como ARIMA (Autoregressive Integrated Moving Average) y modelos de Machine Learning para predecir valores futuros de la serie.

Etapas Preliminares del Análisis

Etapas Preliminares del Análisis

Antes de iniciar el análisis de una serie de tiempo, es esencial llevar a cabo una serie de etapas preliminares que establecen las bases para un estudio efectivo:

- **Experticia en el Ámbito de Aplicación.**
- **Definición de Objetivos.**
- **Recolección de Datos.**
- **Análisis Exploratorio.**
- **Representación de Componentes.**

Etapas Preliminares del Análisis

Definición de Objetivos

Es crucial definir claramente los objetivos del análisis. Esto incluye preguntas como:

- ¿Qué aspectos específicos de la serie se van a analizar?
- ¿Cuáles son los objetivos de dicho análisis? (por ejemplo, entender patrones, predecir tendencias, identificar causas)
- ¿Se enfocará en una parte particular de la serie o en su totalidad?
- ¿Se busca analizar picos, valles u otros patrones específicos?

Establecer los objetivos de manera clara ayudará a definir el alcance del análisis.

Etapas Preliminares del Análisis

Recolección de Datos

Una vez definidos los objetivos, es necesario determinar la frecuencia temporal y la granularidad de los datos que se deben recopilar. Esto incluye decidir la periodicidad de adquisición de datos (por ejemplo, diaria, mensual, anual) y el nivel de cobertura geográfica o temática. Estos aspectos dependen directamente de los objetivos del análisis.

Etapas Preliminares del Análisis

Análisis Exploratorio.

El análisis exploratorio es una fase clave que involucra la comprensión inicial de la serie de tiempo. Las herramientas comunes incluyen:

- Visualización de la serie a través de gráficas de tiempo para identificar patrones, periodicidades y problemas potenciales.
- Determinación de componentes sistemáticos y no sistemáticos de la serie, como la tendencia, la estacionalidad y el ruido.

Este análisis ayuda a familiarizarse con los datos antes de aplicar métodos más avanzados.

Etapas Preliminares del Análisis

Representación de Componentes

Dependiendo de los objetivos del análisis, los componentes sistemáticos de la serie de tiempo pueden representarse de manera aditiva o multiplicativa. Esto implica considerar si la serie se descompone en la suma o la multiplicación de dichos componentes, lo cual puede afectar el enfoque del análisis.

Estacionariedad y Función de Autocorrelación

Muchos métodos estadísticos para el análisis de series de tiempo requieren, como paso previo, verificar la estacionariedad de la señal. La estacionariedad se refiere a que las propiedades estadísticas de la serie se mantienen constantes a lo largo del tiempo.

La función de autocorrelación es una herramienta esencial en este contexto, ya que permite comparar la señal con ella misma y revelar información valiosa sobre la estructura temporal de la serie.

Pre-procesamiento de la Serie de Tiempo

- ① Manejo de Datos Faltantes:**
- ② Uniformización del Espaciamiento Temporal.**
- ③ Manejo de Valores Extremos.**
- ④ Elección del Horizonte de Tiempo.**
- ⑤ Partición de los Datos (Solo para Análisis Predictivo).**

El pre-procesamiento de la serie de tiempo es esencial para garantizar que los datos estén listos para el análisis posterior y para obtener resultados confiables y significativos en función de los objetivos establecidos.

Modelo de Aprendizaje de Máquina

Modelo vs Algoritmo

- **¿Qué es un modelo?** *Un modelo es cualquier tipo de función que tiene capacidad predictiva.*
- **¿Qué es un algoritmo?** *Un algoritmo es un conjunto de pasos realizados en orden con el objetivo de realizar una tarea o resolver un problema específico.*
- Un **Modelo de Aprendizaje de Máquina** es el resultado del proceso de entrenar un algoritmo con un conjunto de datos.

Modelo de Aprendizaje de Máquina

“Todos los modelos son erróneos, pero algunos son útiles.”

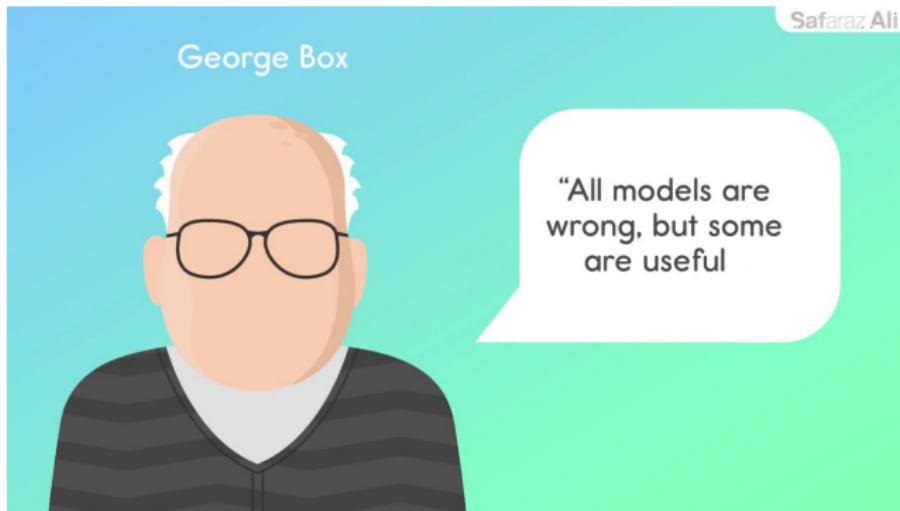


Figure 3: <https://www.youtube.com/@SafarazAli>

Modelo de Aprendizaje de Máquina

“La verdad es demasiado complicada como para permitir algo más que una aproximación.”



Figure 4: <https://www.youtube.com/@SafarazAli>

En Resumen

- El **algoritmo** es como las instrucciones para construir una máquina de predicción.
- El **modelo** es la máquina ya construida y entrenada para realizar una tarea específica.

Modelos de Aprendizaje de Máquina

El Problema

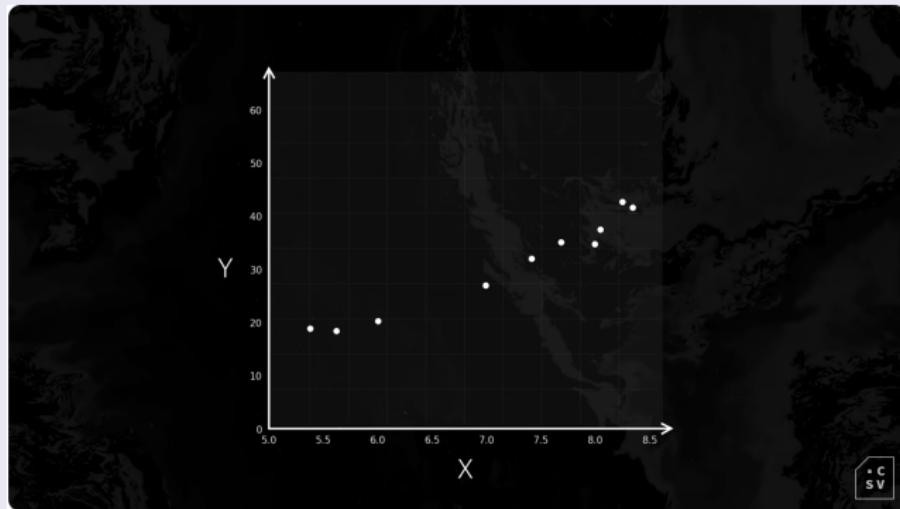


Figure 5: <https://www.youtube.com/>

Modelos de Aprendizaje de Máquina

El Modelo

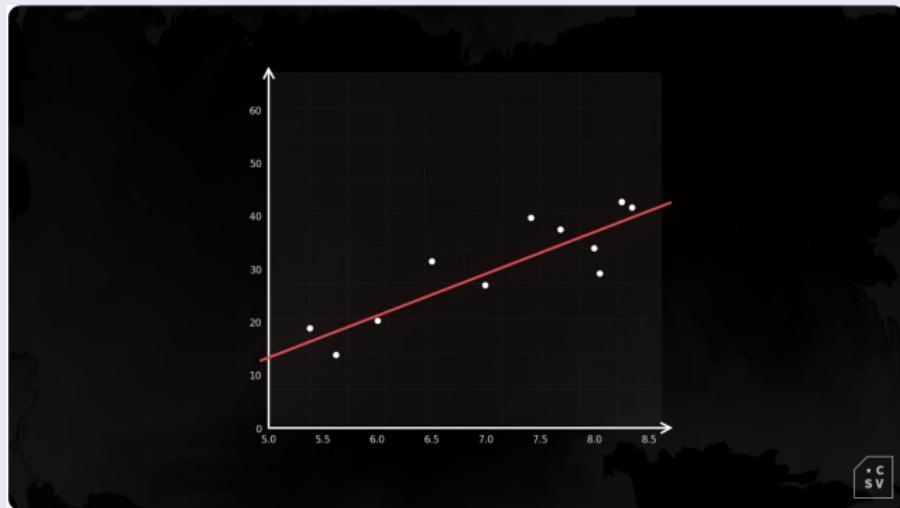


Figure 6: <https://www.youtube.com/>

Modelos de Aprendizaje de Máquina

El Algoritmo

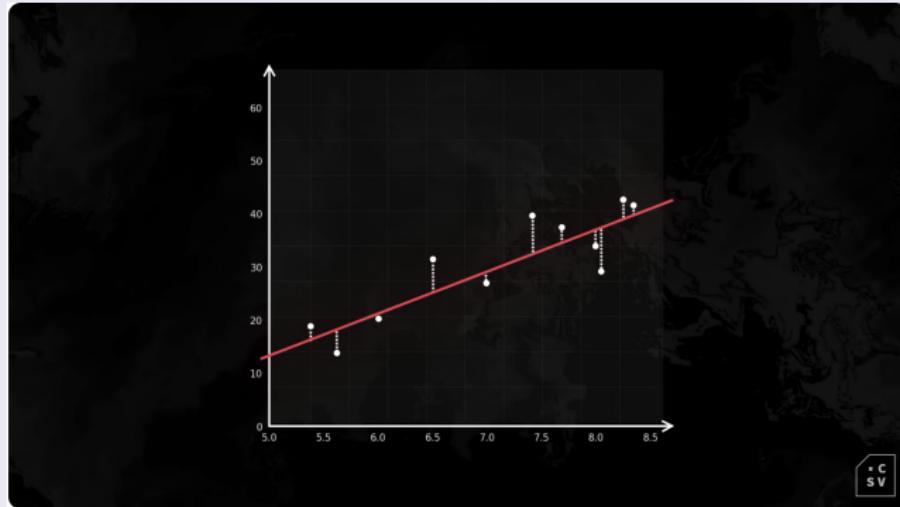


Figure 7: <https://www.youtube.com/>

El Conjunto de Datos

- **Características (Features):**

- Representan las entradas o datos de entrada en un modelo de Machine Learning.
- Son las variables o atributos que se utilizan para hacer predicciones.
- Ejemplos: píxeles de una imagen, palabras en un documento, mediciones en un conjunto de datos.

- **Etiquetas (Labels):**

- Representan las salidas o resultados deseados en un modelo de Machine Learning.
- Son las respuestas que el modelo debe aprender a predecir.
- Ejemplos: categorías de imágenes (perro, gato), clasificación de texto (positivo, negativo), valor numérico de una propiedad.

- **Ejemplos (Examples):** son instancias de datos utilizadas para entrenar y evaluar modelos de aprendizaje automático.

Conceptos en Aprendizaje de Máquina

FEATURES						LABELS
EXAMPLES	Runner ID	Run	Runner time	Elevation	FUN	
AV3DE X8KGF BH9IU	AV3DE	Boston Marathon	03:40:32	1,300ft	Low	High
	X8KGF	Seattle Oktoberfest 5k	00:35:40	0ft	High	
	BH9IU	Houston half-marathon	02:01:18	200ft	Medium	

Figure 8: <https://pair.withgoogle.com/chapter/data-collection/>

El dataframe anterior contiene datos sobre carreras que una aplicación podría usar para entrenar un modelo de aprendizaje automático para predecir qué tan divertida será una carrera determinada.

Conceptos en Aprendizaje de Máquina

Datos del Titanic

passenger	survived	name	sex	age	sibsp	parch	ticket	fare	cabin	embarked	boat	body	home.dest
1	1	Allison, Miss. Elisabeth Walton	female	29	0	0	24160	121,3375	B5	5	2		St Louis, MD
1	1	Allison, Master. Hudson Trevor	male	0.9167	1	2	113781	151,5500	C22 C26 S	11			Montreal, PQ / Chesterville, ON
1	0	Allison, Miss. Helen Loraine	female	2	1	2	113781	151,5500	C22 C26 S				Montreal, PQ / Chesterville, ON
1	0	Allison, Mr. Hudson Joshua Creighton	male	30	1	2	113781	151,5500	C22 C26 S	135			Montreal, PQ / Chesterville, ON
1	0	Allison, Mrs. Hudson J C (Bessie Waldo Daniels)	female	25	1	2	113781	151,5500	C22 C26 S				Montreal, PQ / Chesterville, ON
1	1	Anderson, Mr. Harry	male	48	0	0	19952	30,6500	E12	S	3		New York, NY
1	1	Andrews, Miss. Korrelle Theodosia	female	63	1	0	13902	77,9583	O7	S	10		Hudson, NY
1	0	Andrews, Mr. Thomas J	male	39	0	0	110950	100,0000	A86	S			Belfast, N.I.
1	1	Apeltor, Mrs. Edward Dale (Charlotte Lamson)	female	53	2	0	11769	51,4792	C101	S	0		Bayside, Queens, NY
1	0	Artigaveyela, Mr. Ramon	male	71	0	0	11694	85,042	C				Montevideo, Uruguay
1	0	Astor, Col. John Jacob	male	47	1	0	PC 17757	227,5250	C52 C54 C	124			New York, NY
1	1	Astor, Mrs. John Jacob (Madeleine Talbridge Force)	female	38	1	0	PC 17757	227,5250	C52 C54 C	4			New York, NY
1	1	Aubart, Mme. Leonine Pauline	female	24	0	0	PC 17477	69,3000	B35	C	9		Paris, France
1	1	Barber, Miss. Ellen "Nellie"	female	36	0	0	19877	78,8500	S	G			
1	1	Barkworth, Mr. Algernon Henry Wilson	male	80	0	0	27042	30,0000	A23	S	B		Hessle, Yorks
1	0	Bassanezi, Mr. John D	male	0	0	0	PC 17313	25,9500	S				New York, NY
1	0	Baxter, Mr. Quigg Edmond	male	24	0	1	PC 17558	247,5106	B58 B60 C				Montreal, PQ
1	1	Baxter, Mrs. James (Helene DeLaudeniere Chaput)	female	50	0	1	PC 17558	247,5106	B58 B60 C	6			Montreal, PQ
1	1	Bazzani, Miss. Alberta	female	32	0	0	11813	76,2937	D15	C	8		
1	0	Beattie, Mr. Thomson	male	36	0	0	13050	75,2437	C6	C	A		Winnipeg, MB
1	1	Beccwith, Mr. Richard Leonard	male	37	1	1	11751	52,5542	D35	S	5		New York, NY
1	1	Beccwith, Mrs. Richard Leonard (Sallie Monpony)	female	47	1	1	11751	52,5542	D35	S	5		New York, NY
1	1	Behr, Mr. Karl Howell	male	26	0	0	111369	30,0000	C148	C	5		New York, NY
1	0	Biddu, Miss. Rosalie	female	42	0	0	PC 17557	227,5250	C	4			

Datos de entrenamiento



x_1

y

Variable dependiente

x_2, \dots, x_{13}

Variables independientes

Datos de test



Figure 9: <https://es.slideshare.net/JavierEsteveMeli/>

iprea@udistrital.edu.co

Series de Tiempo con Bosques Aleatorios (Ra

2 Diciembre de 2023

26 / 113

Conceptos en Aprendizaje de Máquina

Categorías

- **Machine Learning Supervisado:**

- Entrenado con datos que especifican tanto la entrada como la salida.
- Por ejemplo, imágenes de números escritos a mano etiquetados con los números correspondientes.
- Reconoce patrones y relaciones entre las características y las etiquetas conocidas.

- **Machine Learning No Supervisado:**

- Entrenado con datos sin etiquetar.
- Analiza datos para encontrar patrones o conexiones significativas.
- Agrupa datos en categorías o encuentra estructuras ocultas.
- Ejemplo: agrupación de noticias en categorías como deportes o crímenes.

Conceptos en Aprendizaje de Máquina

Categorías principales del Aprendizaje de Máquina

Técnicas de Machine Learning

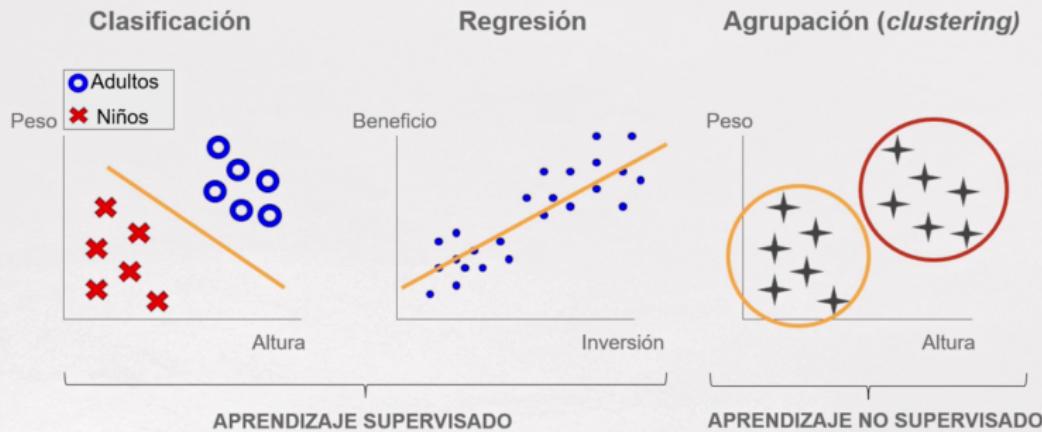


Figure 10: <https://openwebinars.net/blog/modelos-de-machine-learning/>

Tipos de Modelos en Aprendizaje de Máquina

Modelos Supervisados

- Regresión lineal
- Regresión logística
- Máquinas de soporte vectorial (SVM)
- Árboles de decisión
- Redes neuronales

Tipos de Modelos en Aprendizaje de Máquina

Modelos No Supervisados

- K-Means (clustering)
- Análisis de Componentes Principales (PCA)
- Redes Generativas Adversariales (GANs)
- Reglas de asociación
- Detección de anomalías

Los Datos en el Aprendizaje de Máquina

Partición de la Data

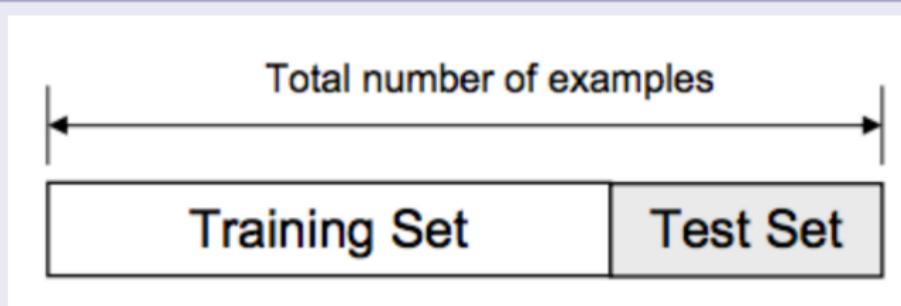


Figure 11: <http://exponentis.es/como-dividir-un-conjunto-de-entrenamiento-en-dos-partes-train-test-split>

Partición de la Data

Partición de la Data

- **Conjunto de Entrenamiento:** Este conjunto se utiliza para entrenar el modelo. Contiene una parte significativa de los datos y es donde el algoritmo de aprendizaje automático "aprende" a partir de las observaciones. El modelo ajusta sus parámetros utilizando este conjunto de datos.
- **Conjunto de Prueba:** Este conjunto se utiliza para evaluar el rendimiento del modelo una vez que ha sido entrenado. Contiene datos que el modelo no ha visto durante el proceso de entrenamiento. Se utiliza para simular cómo el modelo se comportaría en la práctica al hacer predicciones en datos no vistos.

Función de Costo en Modelos

Función de Costo en modelos supervisados

- También conocida como función de pérdida o función de error.
- Mide qué tan bien se hacen las predicciones del modelo en comparación con los valores reales en el conjunto de entrenamiento.

Función de Costo

Función de Costo en Modelos Supervisados

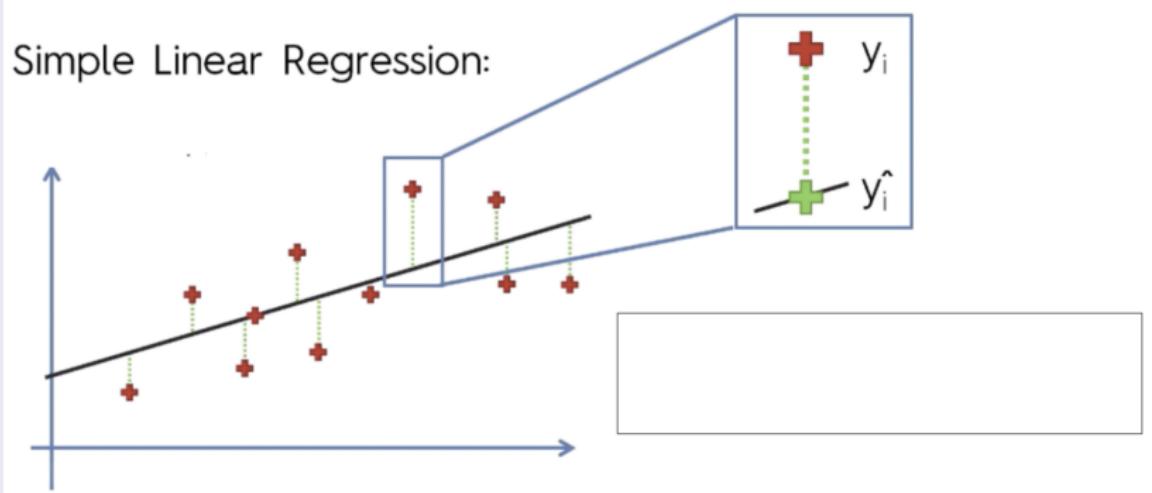


Figure 12: <https://www.themachinelearners.com/>

Función de Costo en Modelos Supervisados

- La función de costo cuantifica la discrepancia entre las predicciones del modelo y los valores reales.
- Evalúa la calidad de las predicciones en el conjunto de entrenamiento.
- Durante el proceso de entrenamiento, la función de costo se utiliza para ajustar los parámetros del modelo.
- El objetivo es encontrar los valores de parámetros que minimizan la función de costo.

Función de Costo en Modelos Supervisados

La elección de la función de costo depende del tipo de problema de Machine Learning:

- **MSE (Mean Squared Error)**: Utilizado en problemas de regresión.
- **Entropía Cruzada (Cross-Entropy)**: Utilizado en problemas de clasificación.

Función de Costo en Modelos No Supervisados

- En modelos no supervisados, también se utilizan funciones de costo.
- Estas funciones miden la calidad de las soluciones o guían el proceso de aprendizaje no supervisado.

Función de Costo

Función de Costo en Modelos No Supervisados

- En K-Medias (Clustering), se utiliza la función de costo “**inercia**” o “**SSE**” para minimizar la distancia de las instancias a los centroides.
- Aunque **PCA** no se ajusta en el sentido tradicional, busca una solución que **optimice la varianza** de las proyecciones de datos.

Árboles de Decisión

Los árboles de decisión son algoritmos versátiles utilizados en tareas de clasificación y regresión. Son especialmente útiles para capturar relaciones no lineales en los datos y son una excelente opción para principiantes en el campo del Machine Learning.

Bootstrap Aggregation (Bagging)

El algoritmo Bagging se utiliza para mejorar la estabilidad y precisión de los modelos, incluyendo árboles de decisión. A través de la combinación de múltiples modelos entrenados con muestras de datos seleccionadas con reemplazo, Bagging reduce la varianza y aumenta la precisión.

Random Forest

Bootstrap Aggregation (Bagging)

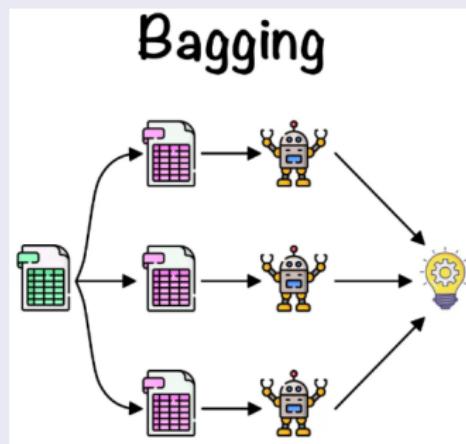


Figure 13: <https://towardsdatascience.com/ensemble-learning-bagging-boosting-3098079e5422>

Random Forest

Bosques Aleatorios (Random Forest)

Bosques Aleatorios son una extensión de Bagging que combina múltiples árboles de decisión mediante la aleatorización de características. Esto produce un modelo robusto y de alto rendimiento que es eficaz en la clasificación y la regresión de series de tiempo.

Random Forest

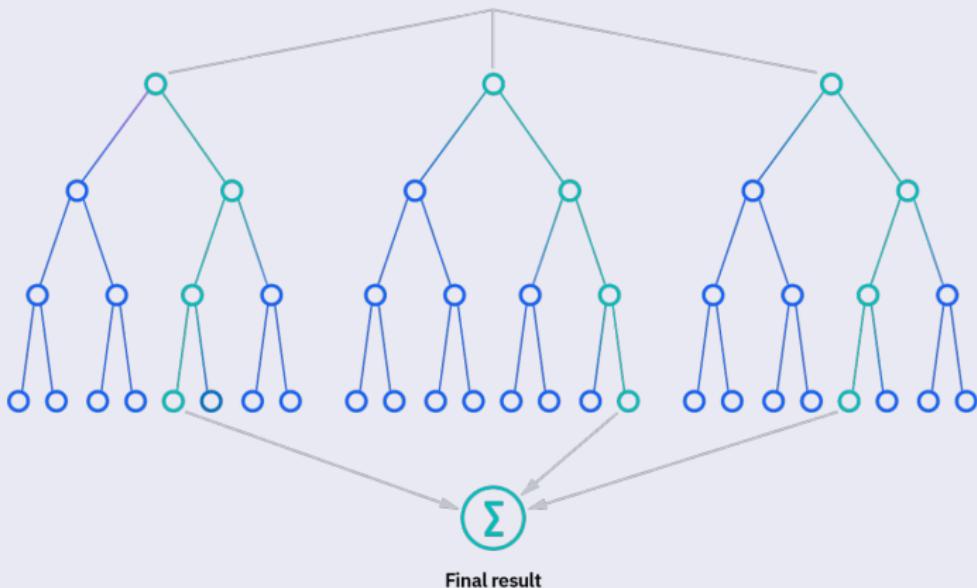


Figure 14: <https://www.themachinelearners.com/>

MODELOS DE APRENDIZAJE DE MÁQUINA

ANÁLISIS DE COMPONENTES PRINCIPALES

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Reducción de dimensionalidad

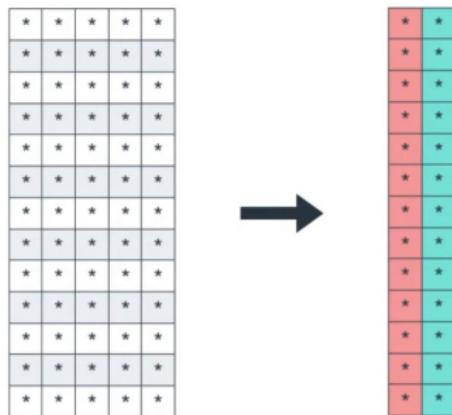


Figure 15: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Tomando una foto



Figure 16: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Tomando una foto



Figure 17: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Tomando una foto



Figure 18: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Tomando una foto



Figure 19: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Reducción de dimensionalidad



Figure 20: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Reducción de dimensionalidad

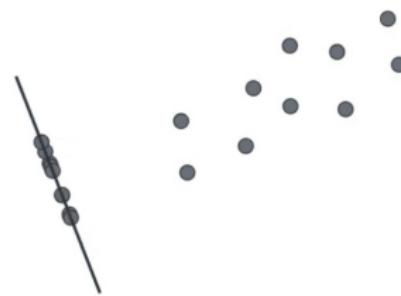


Figure 21: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Reducción de dimensionalidad

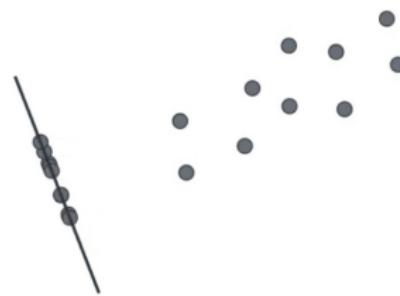


Figure 22: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Reducción de dimensionalidad

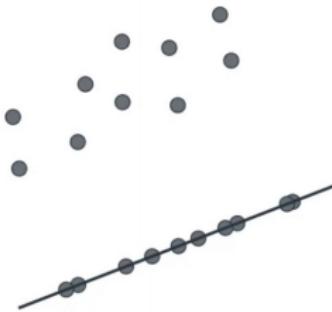


Figure 23: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Reducción de dimensionalidad

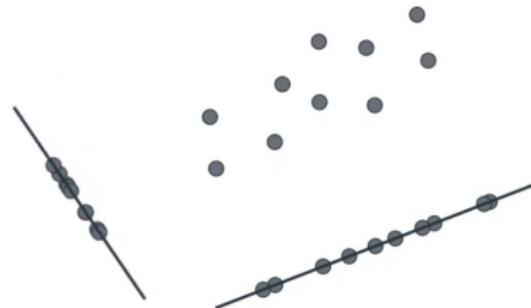


Figure 24: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Reducción de dimensionalidad

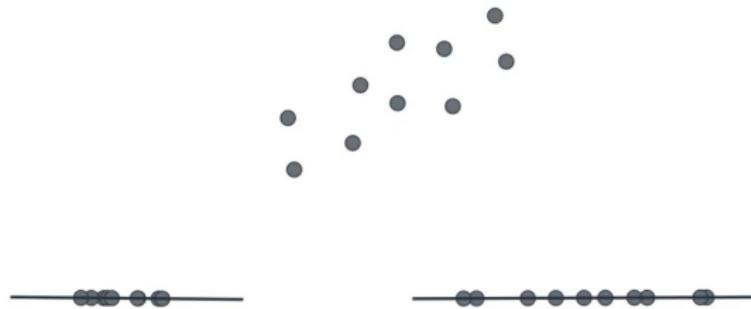


Figure 25: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Datos de casas

Area
Número de habitaciones
Número de baños
Escuelas cercanas
Crimen en el área

Figure 26: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Datos de casas



Figure 27: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

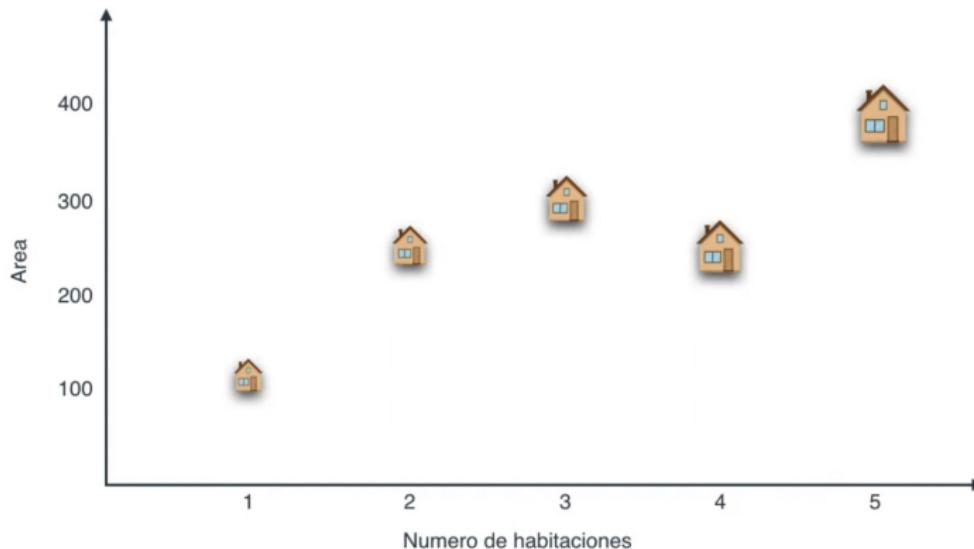


Figure 28: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

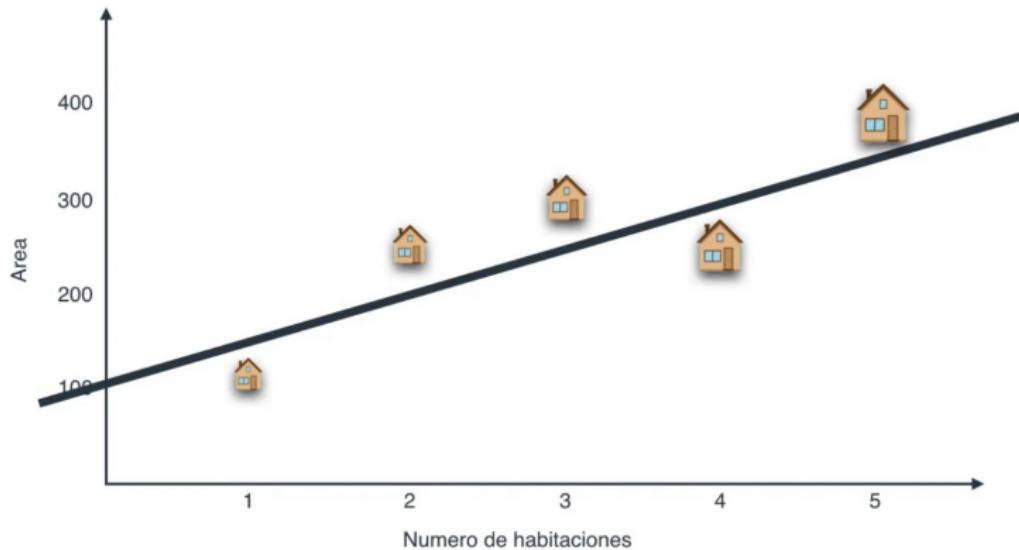


Figure 29: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

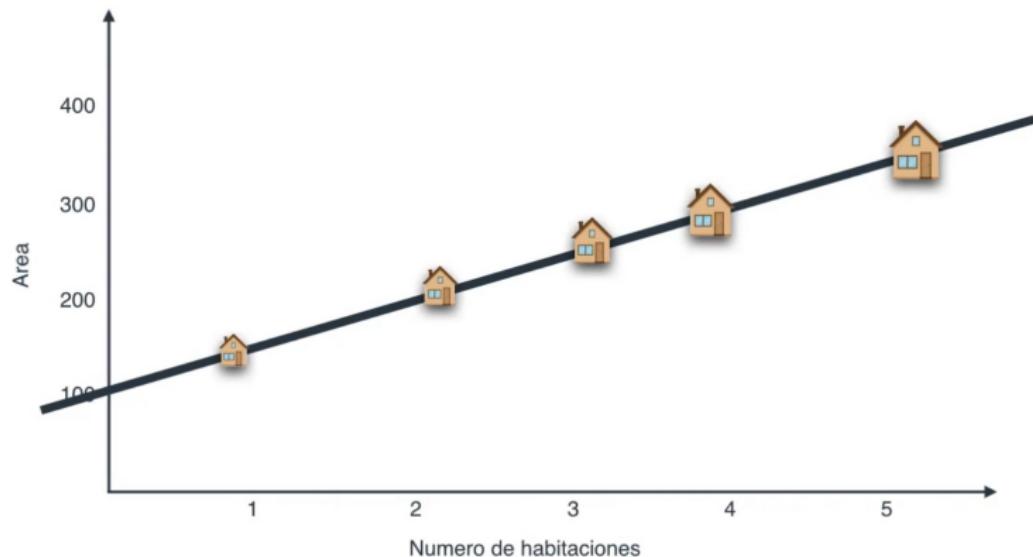


Figure 30: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales



Figure 31: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

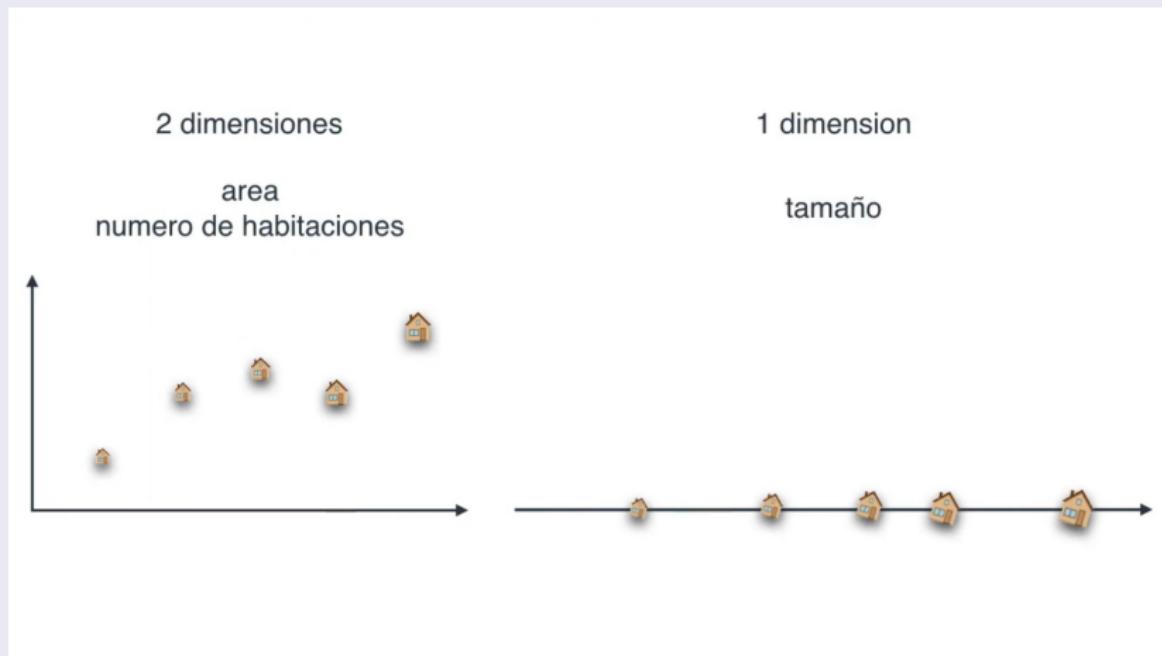


Figure 32: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Datos de casas

5 dimensiones

Area
Número de habitaciones
Número de baños
Escuelas cercanas
crimen en el area

2 dimensions

Tamaño
Ubicación

Figure 33: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Promedio, varianza, covarianza

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Promedio

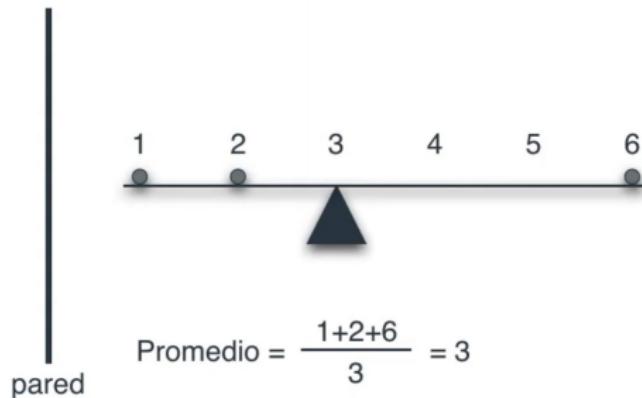


Figure 34: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Varianza



Figure 35: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Varianza

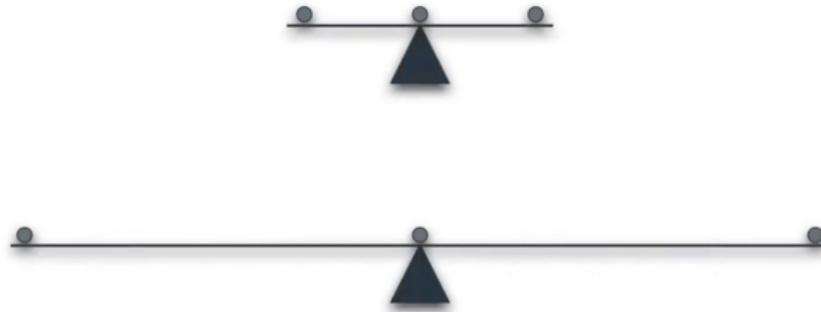


Figure 36: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Varianza



$$\text{Varianza} = \frac{1^2 + 0^2 + 1^2}{3} = 2/3$$



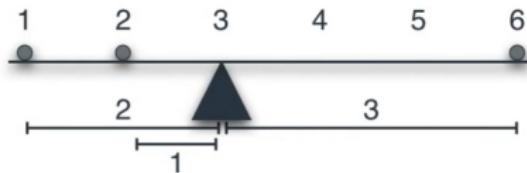
$$\text{Varianza} = \frac{5^2 + 0^2 + 5^2}{3}$$

Figure 37: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Varianza



$$\text{Varianza} = \frac{2^2 + 1^2 + 3^2}{3} = 14/3$$

Figure 38: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Varianza?

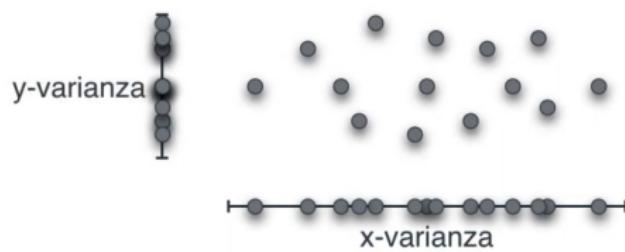
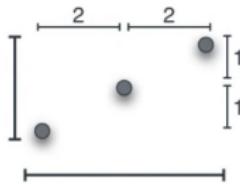
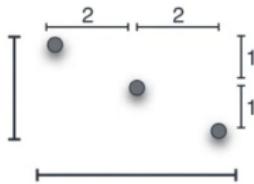


Figure 39: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Varianza?



$$x\text{-varianza} = \frac{2^2+0^2+2^2}{3} = 8/3$$

$$y\text{-varianza} = \frac{1^2+0^2+1^2}{3} = 2/3$$

Figure 40: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

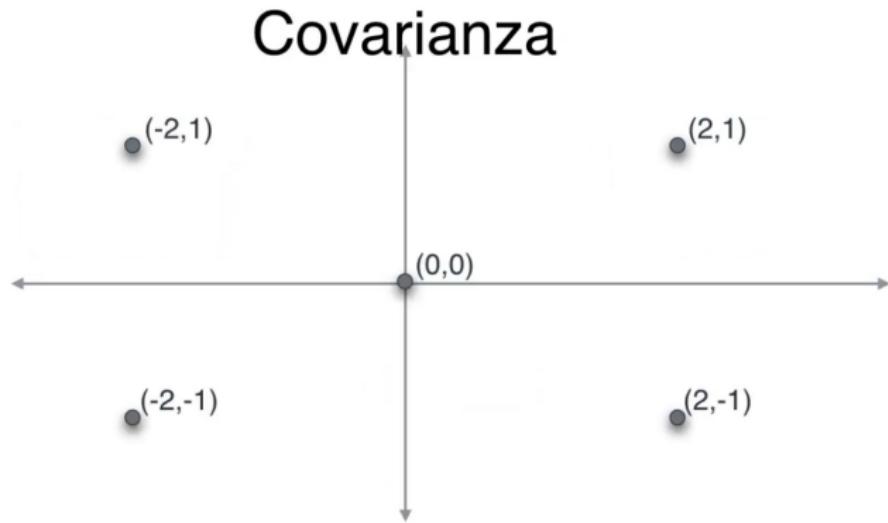


Figure 41: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

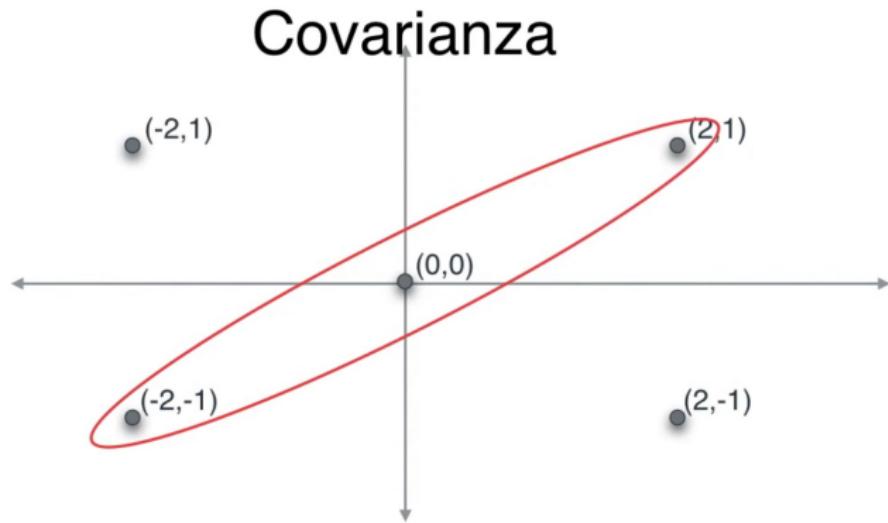


Figure 42: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

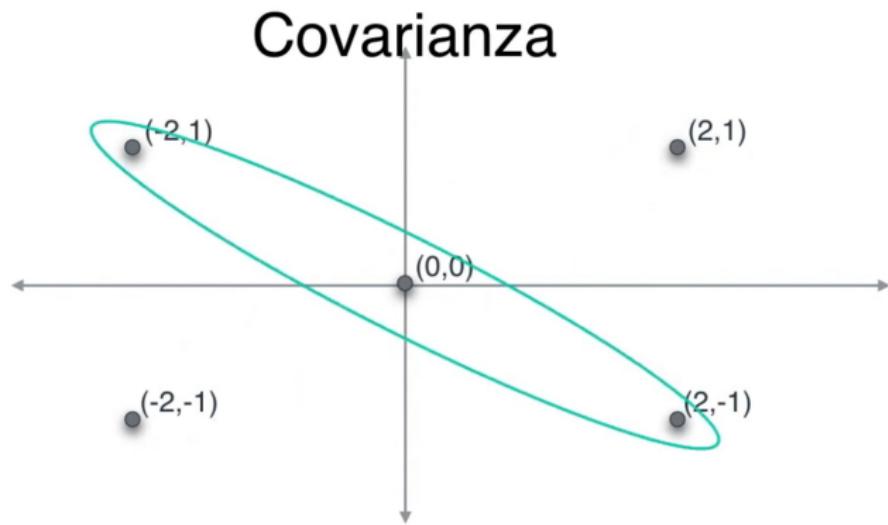


Figure 43: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

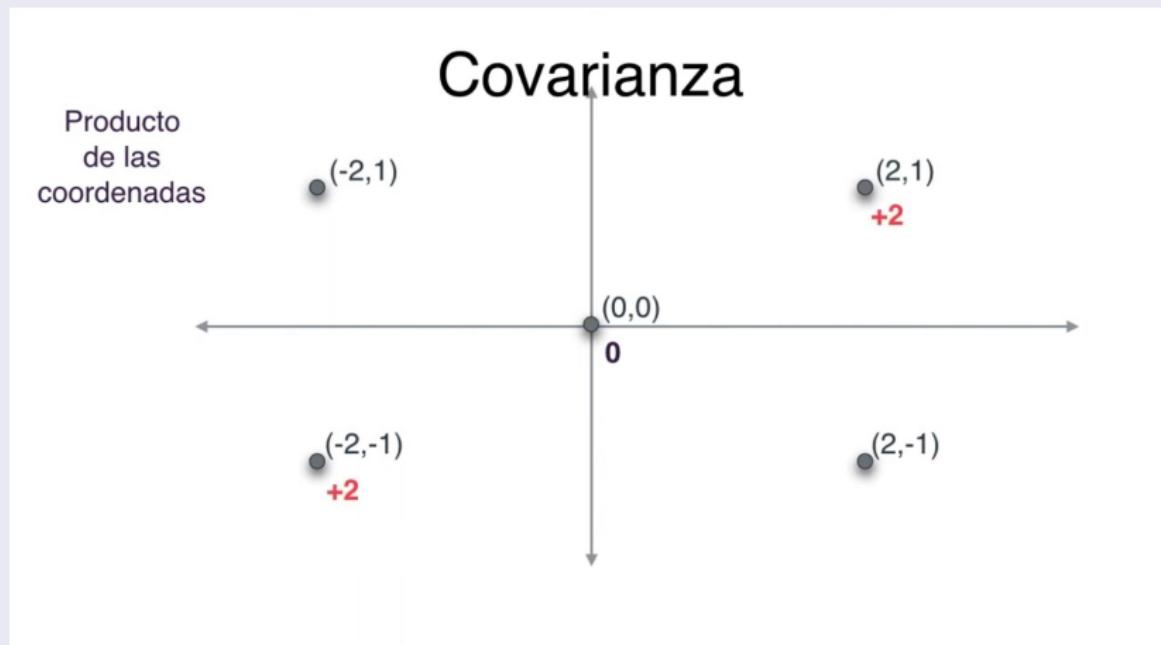


Figure 44: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

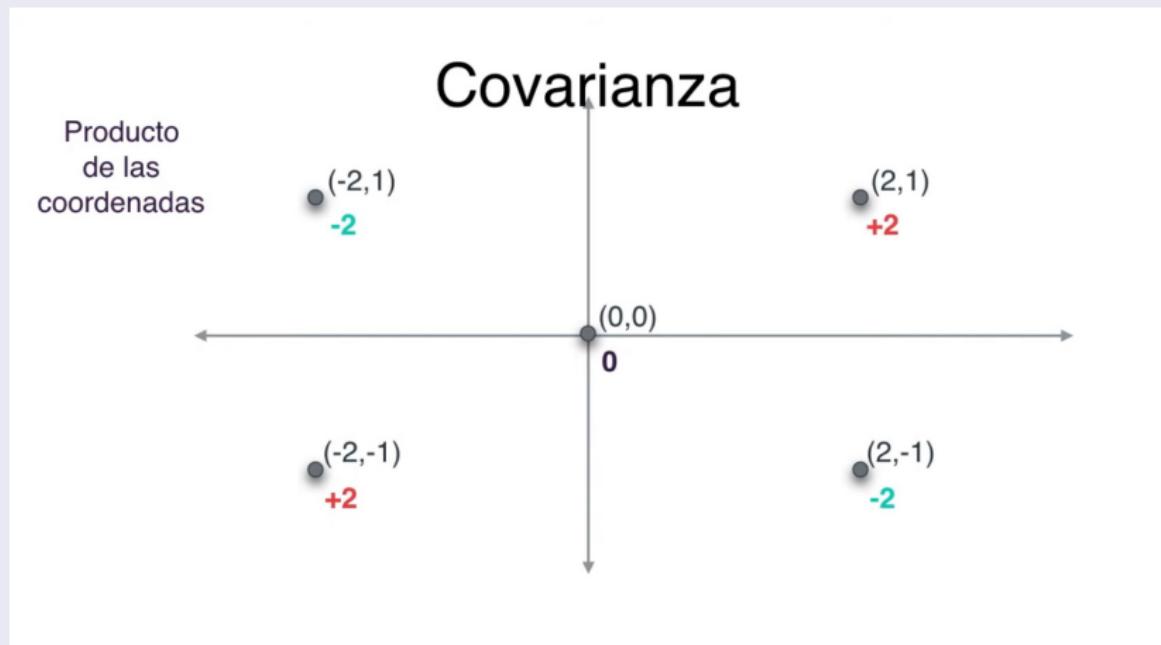
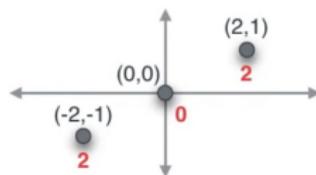
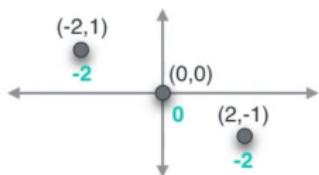


Figure 45: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Covarianza



$$\text{covarianza} = \frac{(-2) + 0 + (-2)}{3} = -4/3$$

$$\text{covarianza} = \frac{2 + 0 + 2}{3} = 4/3$$

Figure 46: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Covarianza

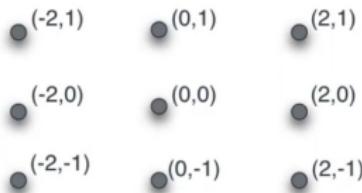
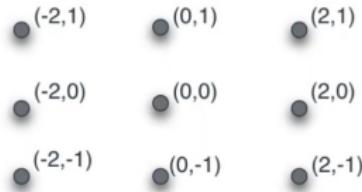


Figure 47: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Covarianza



$$\text{covarianza} = \frac{-2 + 0 + 2 + 0 + 0 + 0 + 2 + 0 + -2}{9} = 0$$

Figure 48: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Covarianza



covarianza
negativa



covarianza cero
(o muy pequeña)



covarianza
positiva

Figure 49: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Valores y vectores propios

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

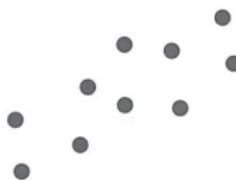


Figure 50: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

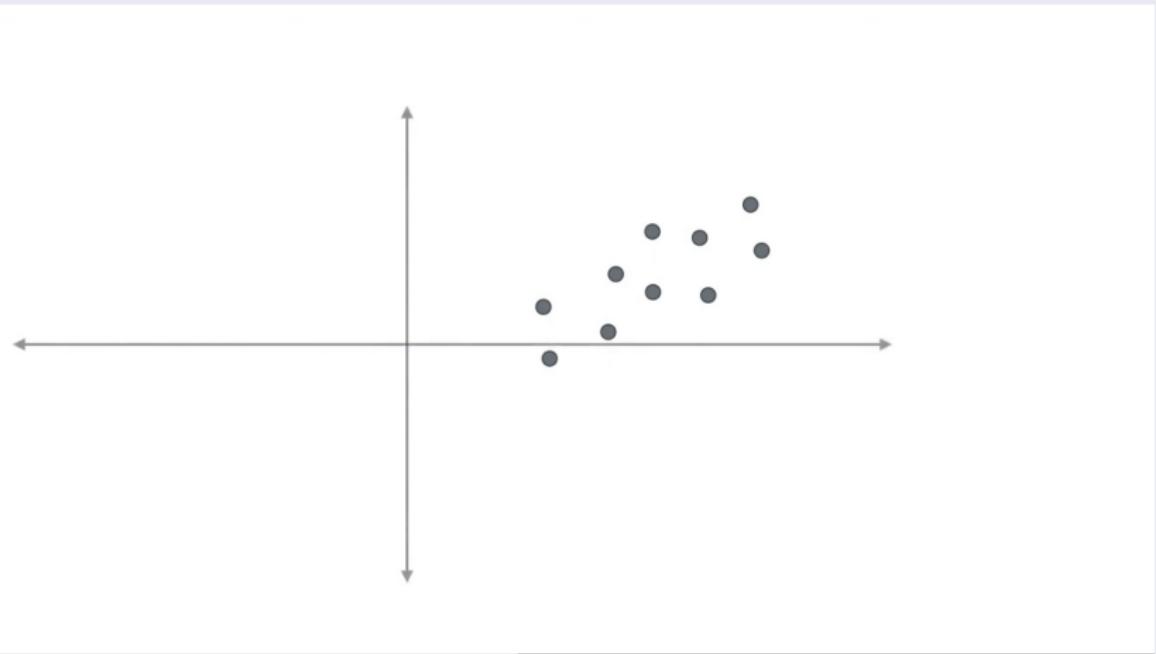


Figure 51: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

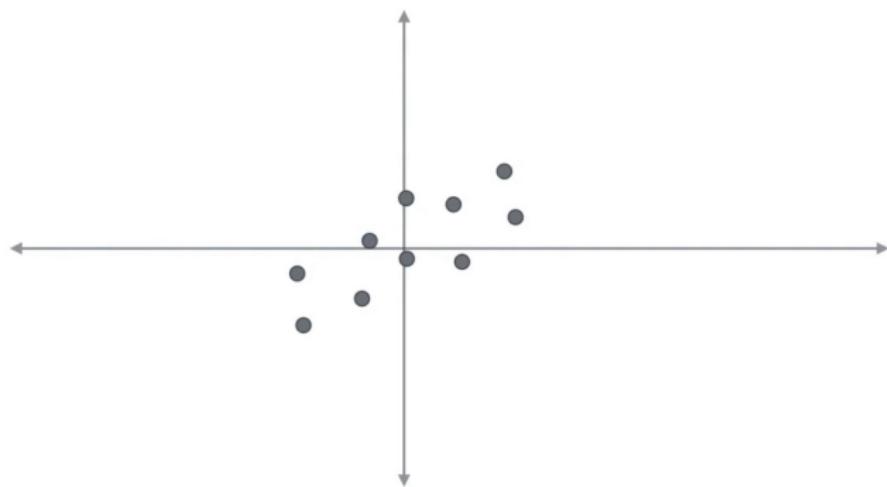


Figure 52: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Matriz de covarianza

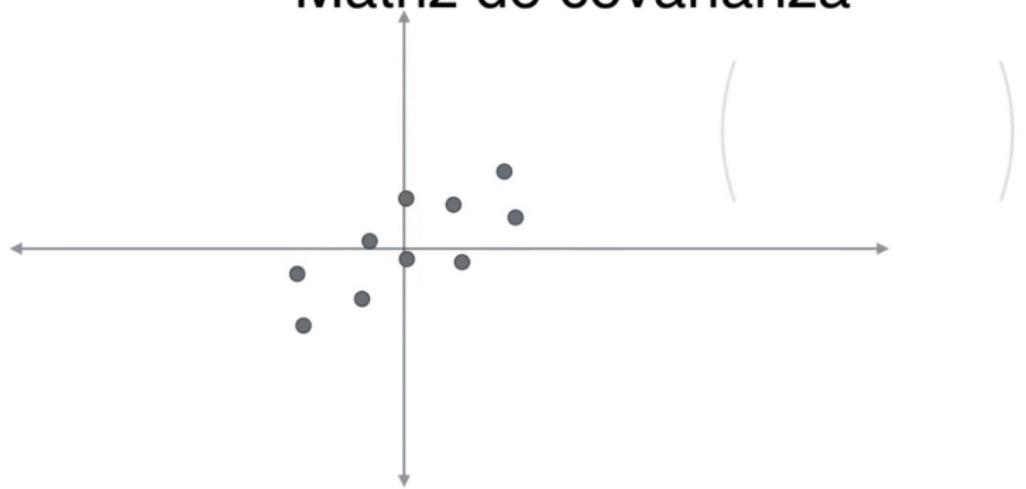


Figure 53: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Matriz de covarianza

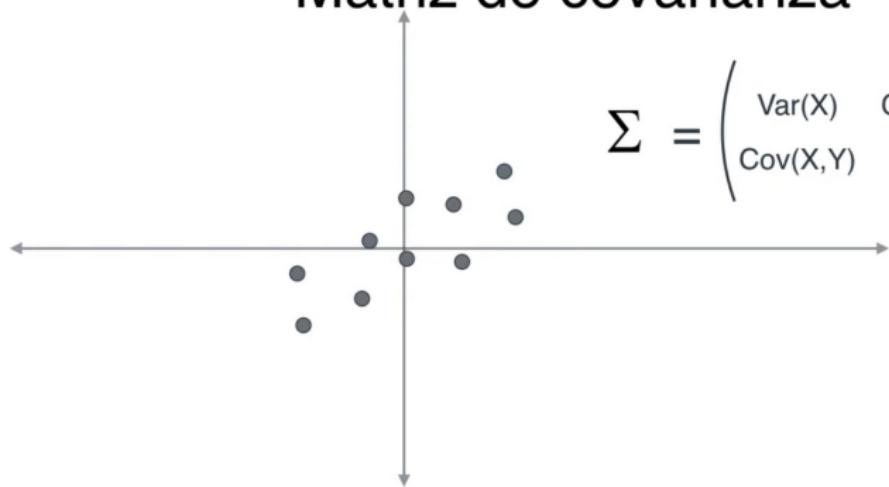


Figure 54: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Matriz de covarianza



$$\Sigma = \begin{pmatrix} \text{Var}(X) & \text{Cov}(X, Y) \\ \text{Cov}(X, Y) & \text{Var}(Y) \end{pmatrix}$$

Figure 55: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Matriz de covarianza

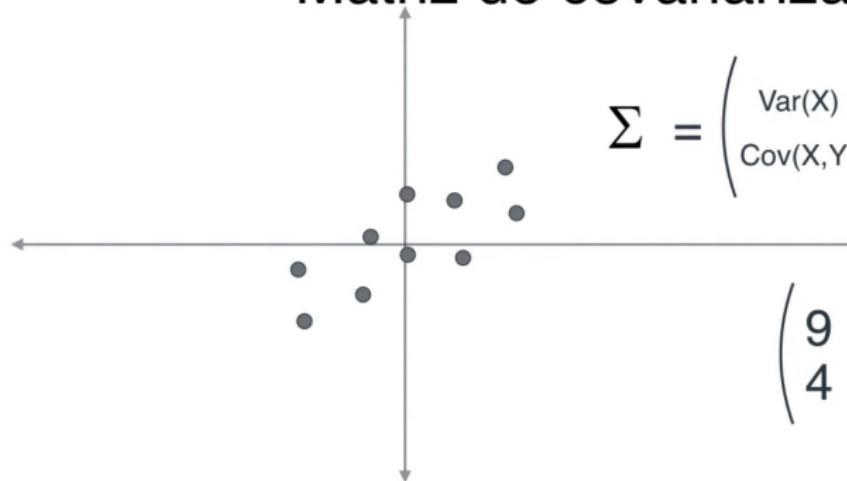


Figure 56: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Transformaciones lineales

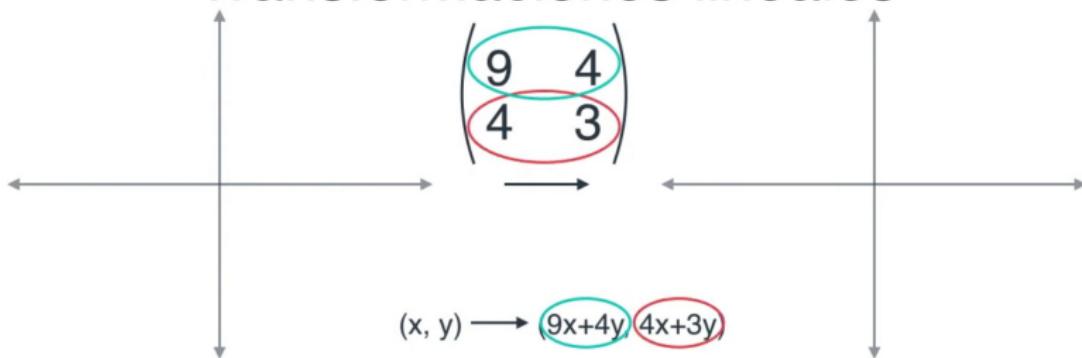


Figure 57: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Transformaciones lineales

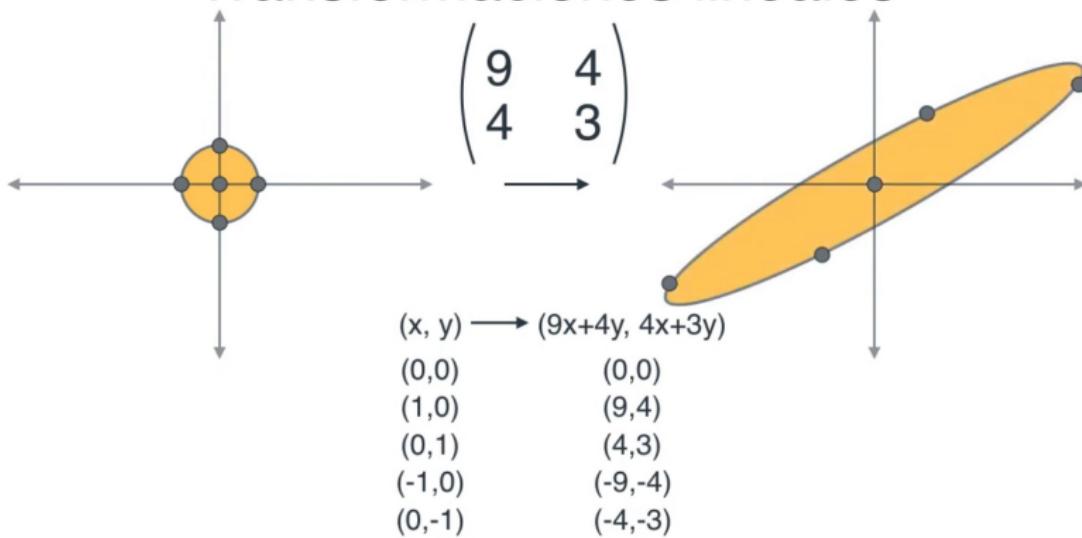


Figure 58: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Transformaciones lineales

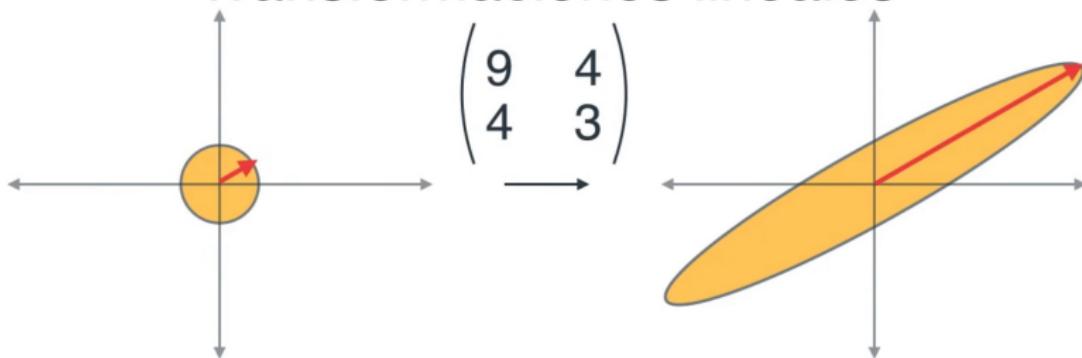


Figure 59: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Transformaciones lineales

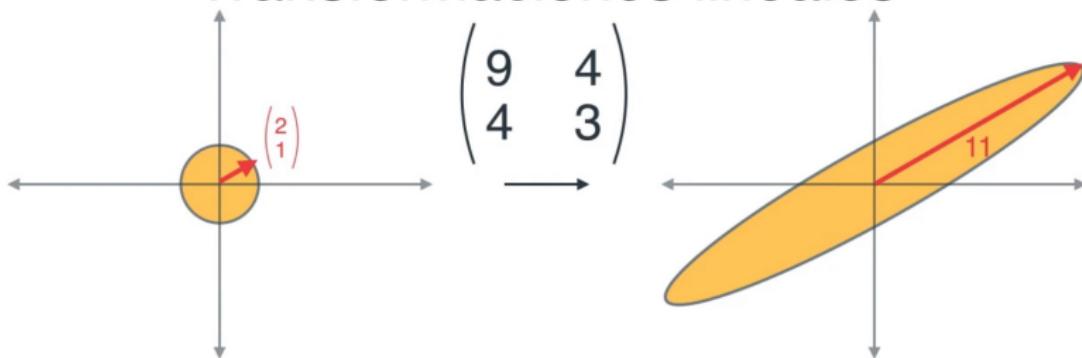


Figure 60: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Transformaciones lineales

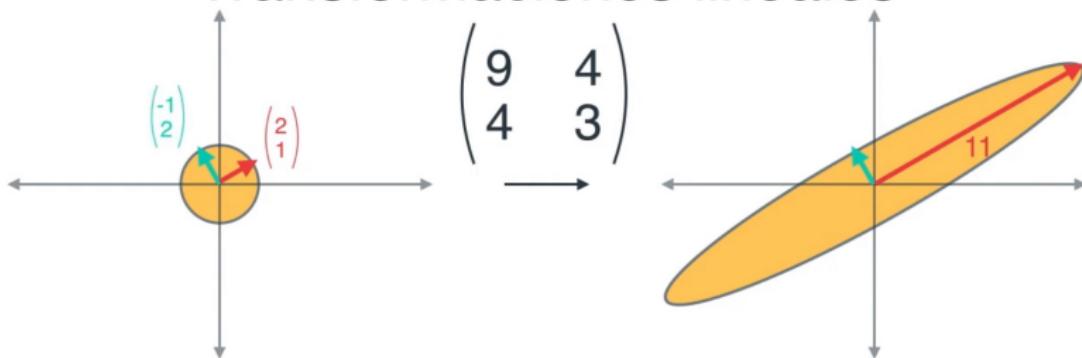


Figure 61: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Transformaciones lineales

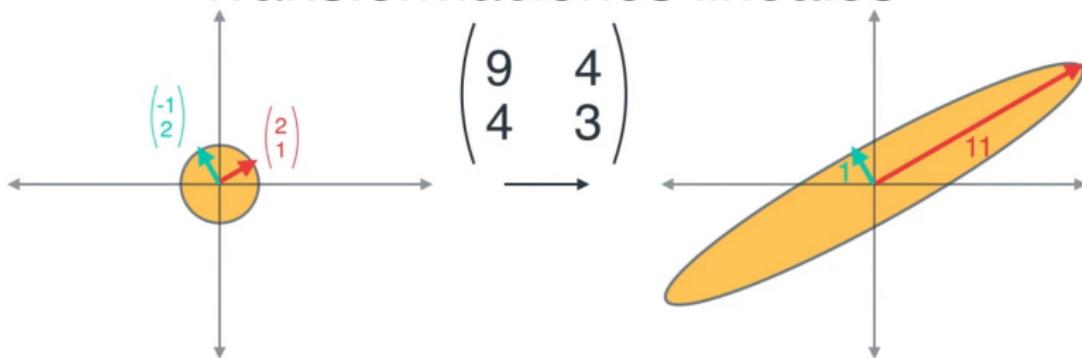


Figure 62: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Transformaciones lineales

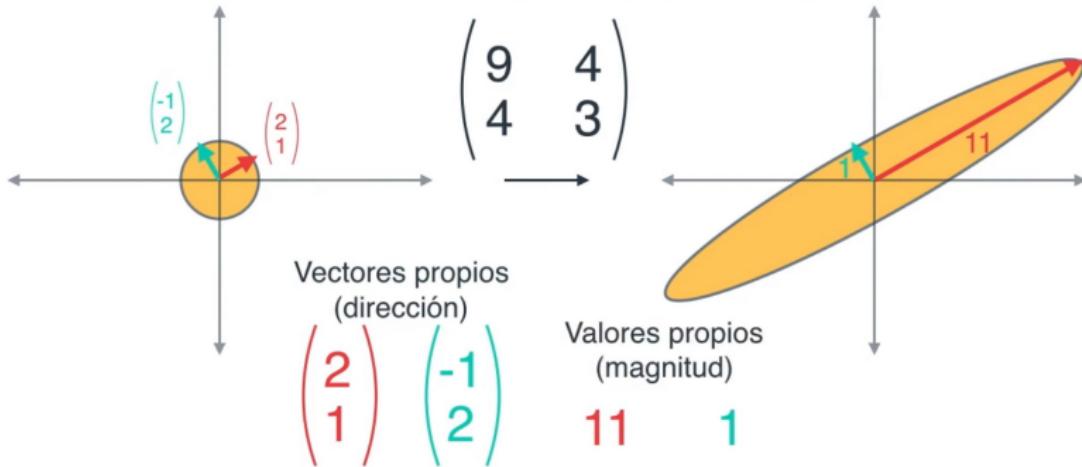


Figure 63: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Transformaciones lineales

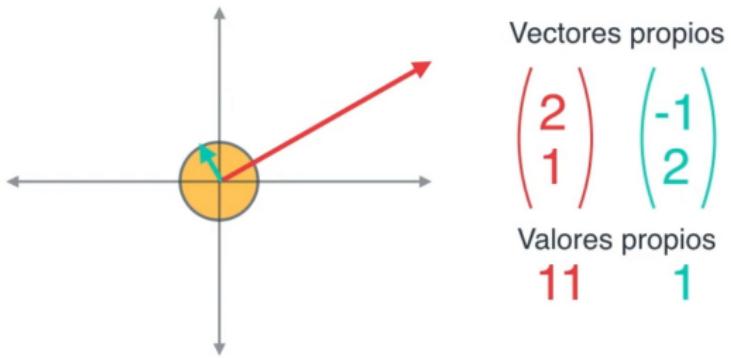


Figure 64: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Transformaciones lineales

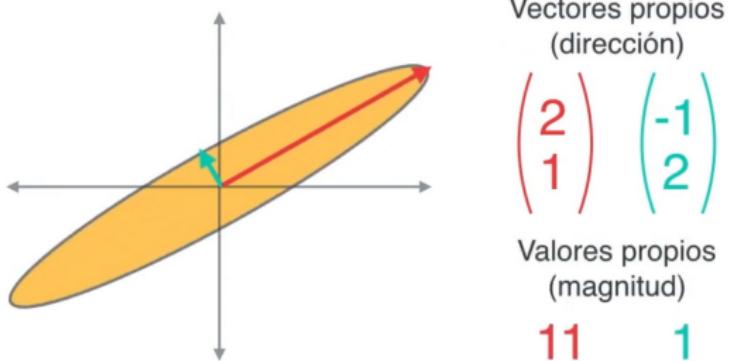


Figure 65: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Valores y vectores propios

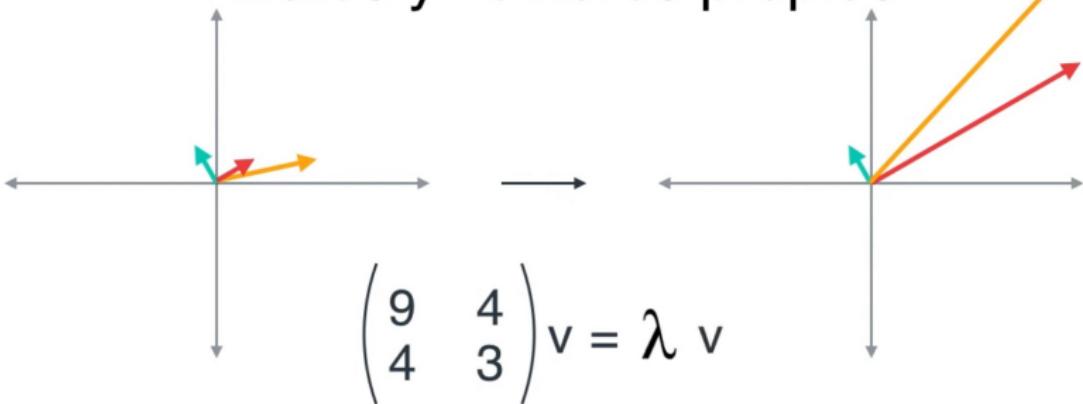


Figure 66: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Valores y vectores propios

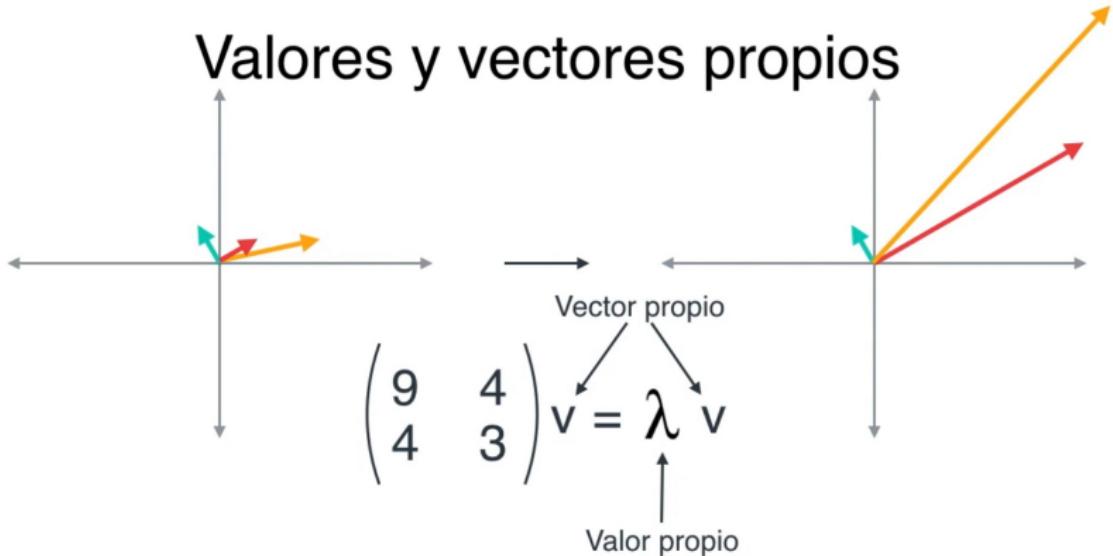


Figure 67: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Análisis de componentes principales (PCA)

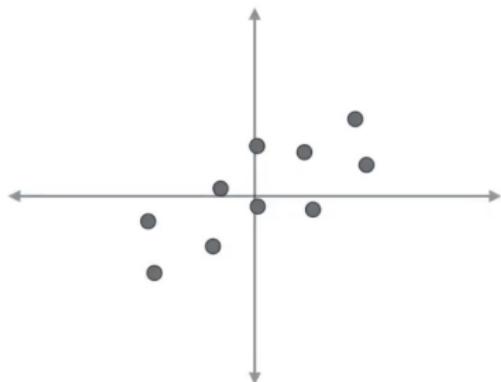
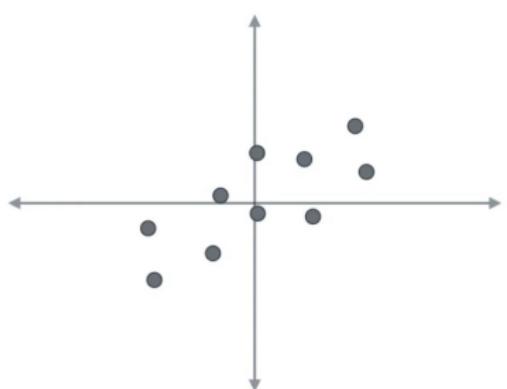


Figure 68: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Análisis de componentes principales (PCA)



$$\Sigma = \begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Vectores propios
(dirección)

$$11$$

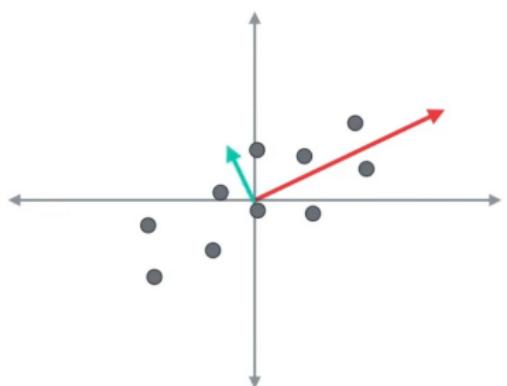
Valores propios
(magnitud)

Figure 69: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Análisis de componentes principales (PCA)



$$\Sigma = \begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Vectores propios
(dirección)

$$11 \quad 1$$

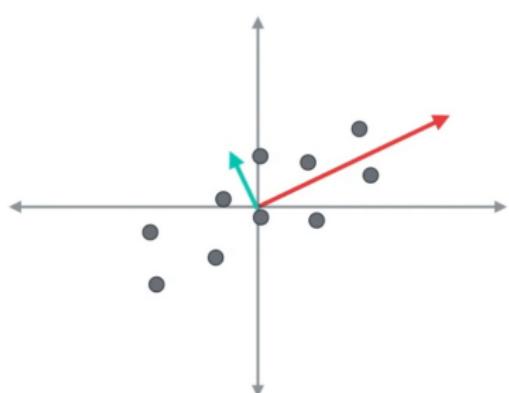
Valores propios
(magnitud)

Figure 70: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Análisis de componentes principales (PCA)



$$\Sigma = \begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Vectores propios (dirección)

11 1 Valores propios (magnitud)

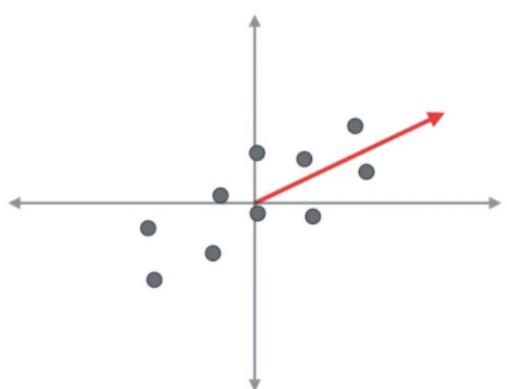


Figure 71: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Análisis de componentes principales (PCA)



$$\Sigma = \begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Vectores propios
(dirección)

$$11$$

Valores propios
(magnitud)

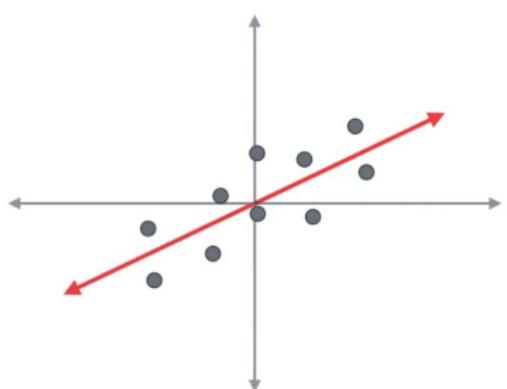


Figure 72: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Análisis de componentes principales (PCA)



$$\Sigma = \begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Vectores propios
(dirección)

11

Valores propios
(magnitud)

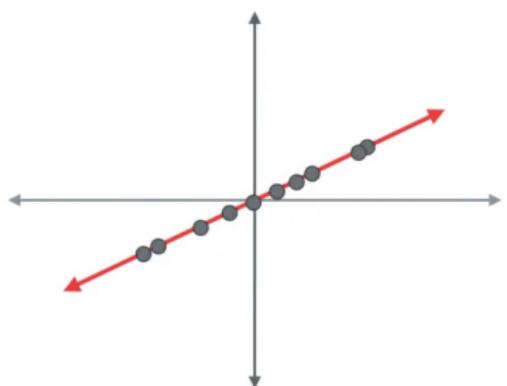


Figure 73: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Análisis de componentes principales (PCA)



$$\Sigma = \begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Vectores propios
(dirección)

$$11$$

Valores propios
(magnitud)

Figure 74: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Análisis de componentes principales (PCA)



Figure 75: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

Análisis de componentes principales (PCA)

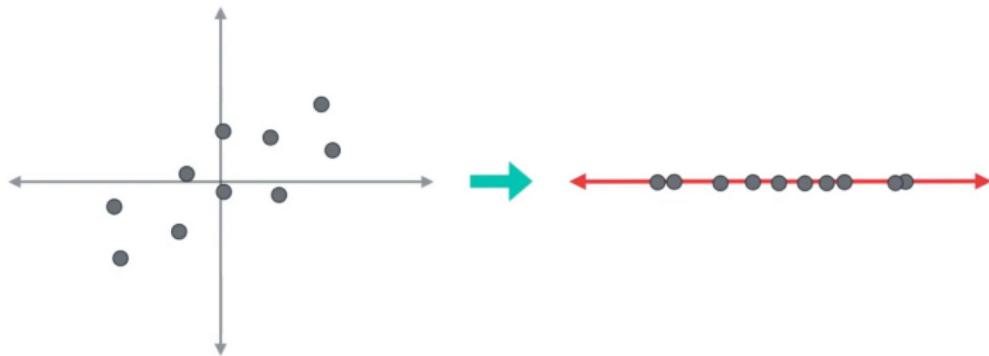


Figure 76: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

PCA

Tabla grande

X1	X2	X3	X4	X5
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*

Matriz de covarianza Vectores y valores propios

$$\begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix} \xrightarrow{\quad} \begin{array}{ll} V_1 & \lambda_1 \\ V_2 & \lambda_2 \\ V_3 & \lambda_3 \\ V_4 & \lambda_4 \\ V_5 & \lambda_5 \end{array}$$

Grande
Pequeño

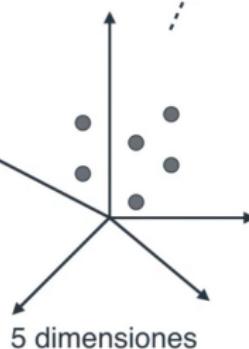


Figure 77: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

PCA

Tabla grande

	X1	X2	X3	X4	X5
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*

Matriz de covarianza Vectores y valores propios

$$\left(\begin{array}{cccc} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{array} \right) \xrightarrow{\quad} \begin{array}{ll} V_1 & \lambda_1 \\ V_2 & \lambda_2 \end{array}$$

Grande
Pequeño

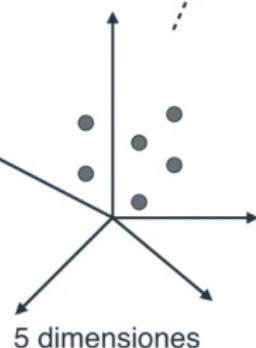


Figure 78: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

PCA

Tabla grande

X1	X2	X3	X4	X5
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*

Matriz de covarianza Vectores y valores propios

$$\begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix} \xrightarrow{\text{---}} \begin{array}{l} V_1 \quad \lambda_1 \\ V_2 \quad \lambda_2 \end{array}$$

Grande

Pequeño

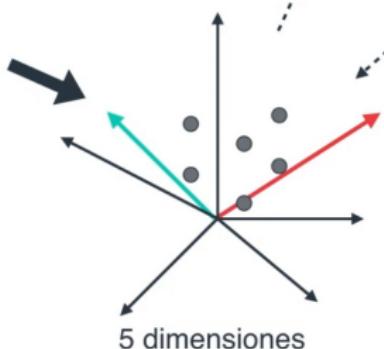


Figure 79: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

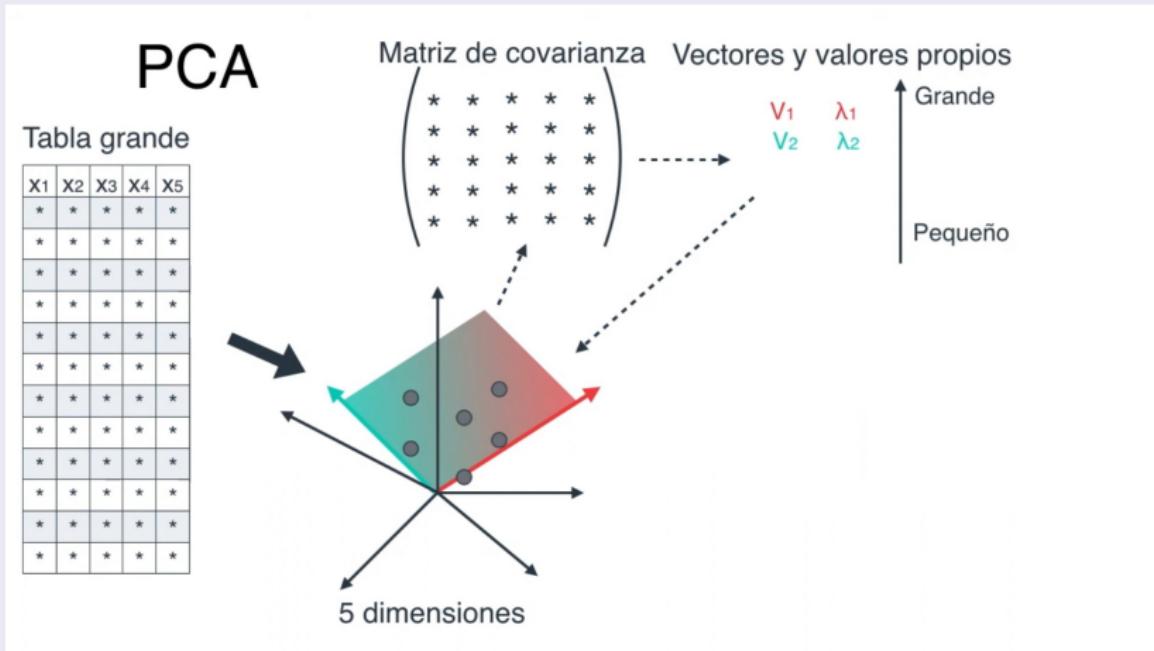


Figure 80: <https://serrano.academy/espanol/>

REDUCCIÓN DE DIMENSIONALIDAD

Análisis de Componentes Principales

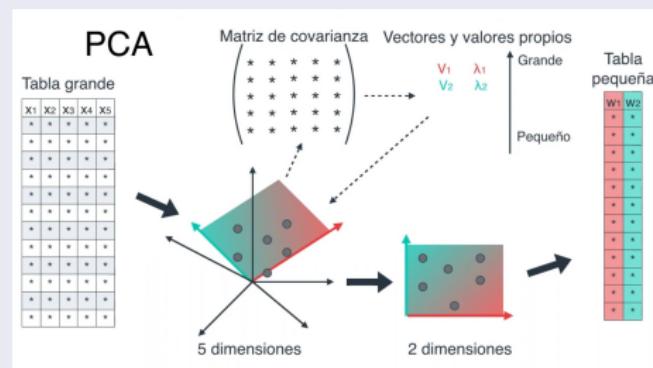


Figure 81: <https://serrano.academy/espanol/>