

Producto Interno de Funciones

Historia y Desarrollo

El concepto de producto interno tiene sus raíces en el desarrollo del álgebra lineal y el análisis funcional. Fue formalizado en el siglo XIX por matemáticos como **Hermann Grassmann** y **David Hilbert**. Grassmann introdujo la idea de un espacio vectorial, mientras que Hilbert desarrolló los espacios que llevan su nombre, donde el producto interno juega un papel crucial.

Definición Matemática

El producto interno de dos funciones f y g en un intervalo $[a, b]$ se define como:

$$\langle f, g \rangle = \int_a^b f(x)g(x) \, dx$$

Este producto interno satisface las siguientes propiedades:

- Linealidad:** $\langle af + bg, h \rangle = a\langle f, h \rangle + b\langle g, h \rangle$
- Simetría:** $\langle f, g \rangle = \langle g, f \rangle$
- Positividad:** $\langle f, f \rangle \geq 0$ y $\langle f, f \rangle = 0$ si y solo si $f = 0$

Conclusión

El producto interno de funciones es una herramienta fundamental en matemáticas, permitiendo definir conceptos como la ortogonalidad y la norma de funciones. Es esencial en el estudio de espacios de Hilbert y tiene aplicaciones en diversas áreas de la ciencia y la ingeniería.

Aplicaciones en Ciencia de Datos

- Análisis de Componentes Principales (PCA):** Utiliza el producto interno para encontrar las direcciones principales de variación en los datos.
- Métodos de Mínimos Cuadrados:** Utiliza el producto interno para minimizar el error en la regresión lineal.
- Procesamiento de Señales:** Utiliza el producto interno para filtrar y analizar señales en diferentes dominios.
- Análisis de Fourier:** Utiliza el producto interno para descomponer funciones en series de senos y cosenos ortogonales.
- Redes Neuronales:** Utiliza el producto interno en la propagación hacia adelante y el cálculo de gradientes.

Ejemplo Práctico: Sistema SCADA en la Industria del Petróleo y Gas

Descripción del Problema

En la industria del petróleo y gas, los sistemas SCADA (Supervisory Control and Data Acquisition) son esenciales para monitorear y controlar operaciones remotas. Un desafío común es la presencia de ruido en los datos de sensores, lo que puede afectar la precisión del monitoreo y control. Nuestro objetivo es filtrar el ruido de los datos de sensores utilizando el complemento ortogonal.

Solución Planteada

Utilizaremos un método basado en la descomposición en valores singulares (SVD) para separar la señal útil del ruido en los datos de sensores. La idea es proyectar los datos en un subespacio generado por los componentes principales y utilizar el complemento ortogonal para eliminar el ruido.

Implementación en Python

Código

```
In [6]: import numpy as np
import matplotlib.pyplot as plt
from sklearn.decomposition import PCA
```

```
In [7]: # Generar datos simulados de sensores
np.random.seed(42)
X = np.random.rand(100, 5) # 100 muestras, 5 características

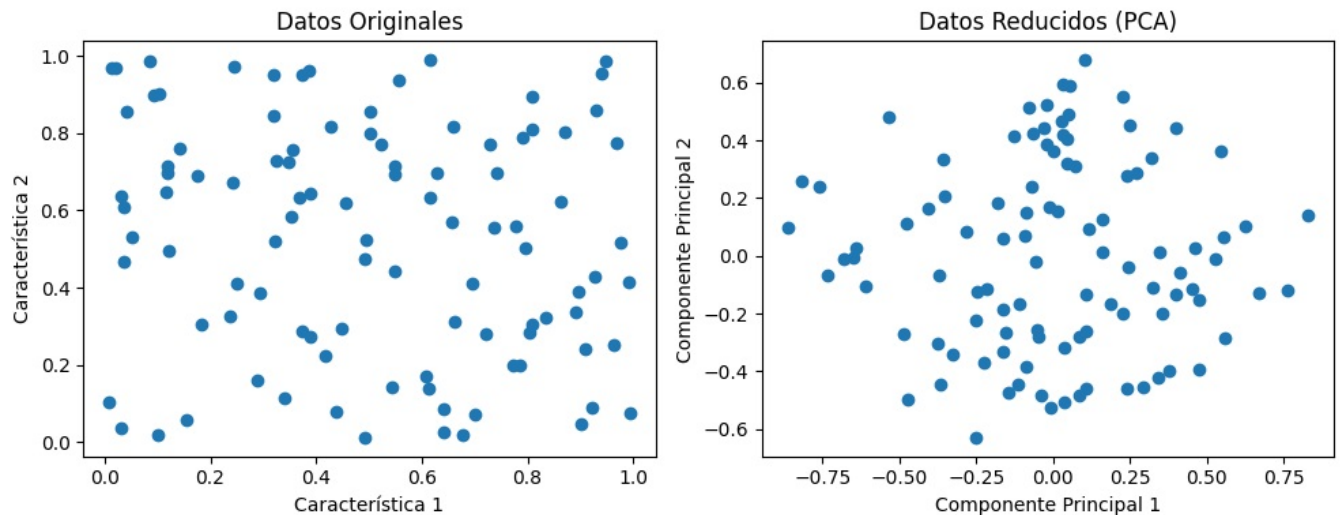
# Aplicar PCA para reducir la dimensionalidad a 2 componentes principales
pca = PCA(n_components=2)
X_pca = pca.fit_transform(X)

# Graficar los datos originales y los datos reducidos
plt.figure(figsize=(10, 4))
```

```
# Datos originales
plt.subplot(1, 2, 1)
plt.scatter(X[:, 0], X[:, 1])
plt.title('Datos Originales')
plt.xlabel('Característica 1')
plt.ylabel('Característica 2')

# Datos reducidos
plt.subplot(1, 2, 2)
plt.scatter(X_pca[:, 0], X_pca[:, 1])
plt.title('Datos Reducidos (PCA)')
plt.xlabel('Componente Principal 1')
plt.ylabel('Componente Principal 2')

plt.tight_layout()
plt.show()
```



4. Uso del Producto Interno de Funciones

Dimension de la dimensionalidad de un dataset, cuando se tienen demasiadas dimensiones este análisis cobra realmente mucha importancia, dado que se seleccionan características fundamentales y permite trabajar de una mejor forma con los datos.

4.1. de esta forma se impactan temas que tiene que ver con el 4.2. overfitting del modelo, 4.3. la alta dimensionalidad y el cómputo disponible, 4.4 además de la redundancia de los datos.

El producto interno de funciones se utiliza aquí para proyectar los datos en el subespacio generado por los componentes principales. Esto permite separar la señal útil del ruido, ya que los componentes principales capturan la mayor parte de la variabilidad de la señal, mientras que el ruido se distribuye en los componentes menores.

Enlaces

<https://medium.com/@roshmitadey/understanding-principal-component-analysis-pca-d4bb40e12d33>

<https://oil-gas.net/implementing-scada-systems-for-enhanced-operational-control/>

http://matematicas.uam.es/~fernando.chamizo/asignaturas/2021alglin/sections/4_3.pdf

<https://mathshistory.st-andrews.ac.uk/Biographies/Grassmann/>

<https://machinelearningmastery.com/article/11751456084/>

<https://programmerclick.com/article/6115655816/>