Instituto de Investigación en Inteligencia Artificial

**Región** Xalapa

Maestría en Inteligencia Artificial

**Análisis discriminante para el caso no lineal**

Protocolo

Presenta:

Ángel García Báez

Director:

Dr. Héctor Gabriel Acosta Mesa

Febrero de 2025

“Lis de Veracruz: Arte, Ciencia, Luz”

Índice

[Índice 3](#_Toc189776959)

[I. Introducción 4](#_Toc189776960)

[1.1 Marco teórico 4](#_Toc189776961)

[1.1.1 Análisis discriminante lineal de Fisher 4](#_Toc189776962)

[1.1.2 Desarrollo matemático 5](#_Toc189776963)

[1.1.3 Escalamiento de los pesos. 6](#_Toc189776964)

[1.1.4 Limitaciones del análisis discriminante lineal 7](#_Toc189776965)

[1.2.1 Expansión de Kronecker 8](#_Toc189776966)

[1.2 Antecedentes 9](#_Toc189776967)

[2 Definición del problema 10](#_Toc189776968)

[3 Propuesta 10](#_Toc189776969)

[4 Justificación 10](#_Toc189776970)

[5 Hipótesis 12](#_Toc189776971)

[6 Objetivos 12](#_Toc189776972)

[6.1 Objetivo principal 12](#_Toc189776973)

[6.2 Objetivos específicos 12](#_Toc189776974)

[7 Metodología 12](#_Toc189776975)

[Referencias 13](#_Toc189776976)

# I. Introducción

## **1.1 Marco teórico**

### 1.1.1 Análisis discriminante lineal de Fisher

El análisis discriminante lineal es una técnica estadística multivariante propuesta por R.A. Fisher en 1936 en su artículo titulado: “The use of multiple measurements in taxonomic problems”.

El análisis discriminante lineal clásico de Fisher trabaja con variables independientes que son de tipo numéricas continuas para poder predecir el resultado de una variable dependiente categórica, donde el caso más sencillo es cuando la variable dependiente es binaría o dicotómica, generando una combinación lineal que tenga la capacidad de poder discriminar o clasificar si una observación pertenece a una u otra clase, con base en sus variables medidas (Lin y Chen, 2004).

Uno de los principales objetivos de la técnica multivariante es generar una regla discriminante (el modelo) que permita clasificar una nueva observación de manera óptima en su clase correspondiente dadas sus características conocidas (Johnson y Wichern, 2007).

La técnica se basa en el cumplimiento de la normalidad multivariante de cada una de las matrices de datos de las clases junto con la homogeneidad de varianzas y covarianzas entre las clases (Silalahi et al., 2016);pese al requerimiento del cumplimiento de estos supuestos, autores como McLachlan (2004) mencionan que la ausencia de estos no presenta graves consecuencias en las reglas discriminantes y en su desempeño obtenido mientras que el desajuste no sea excesivo.

### 1.1.2 Desarrollo matemático

De acuerdo con Johnson y Wichern, 2007:

Sea un conjunto de datos provenientes de una variable aleatoria multivariante:

Donde se tienen observaciones que pertenecen a la clase 1 y observaciones que pertenecen a la clase 2 donde variables independientes. Las respectivas matrices son:

A partir de las matrices de datos de cada clase se obtienen los vectores de medias y matrices de varianzas y covarianzas correspondientes:

Asumiendo que las matrices de varianzas y covarianzas de las 2 clases son estadísticamente iguales entonces se procede a calcular la matriz de varianzas y covarianzas mancomunada de la siguiente forma:

La cual puede ser simplificada como sigue:

Se quiere encontrar un vector que maximice la separación entre grupos como una combinación lineal con la estructura:

donde:

Las proyecciones de las observaciones vienen dadas por:

Se calcula el punto de corte para determinar la asignación que recibirá la observación a una de las dos clases:

Que también se puede expresar de la siguiente manera:

Para la asignación de nuevas observaciones a una de las clases mediante la regla discriminante encontrada, se procede con el siguiente criterio:

Asigna la observación a clase 1 si

Asigna la observación a clase 2 si

### 1.1.3 Escalamiento de los pesos.

El vector de coeficientes calculado puede verse también de la forma , donde es un valor escalar que multiplica al vector de pesos, de tal modo que cualquier valor de hace que los pesos funcionen como regla discriminante.

Con la finalidad de mejorar la interpretabilidad del vector de pesos se recurre a “normalizar” o “escalar” sus elementos. Algunas de las formas más comunes de hacer esto son las siguientes:

Normalización 1: , de esta manera el vector consigue tener longitud unitaria.

Normalización 2: , de esta manera el primer elemento del nuevo vector de coeficientes es 1.

Para la normalización 1, el valor escalar es y para la normalización 2, .

La normalización de los pesos es recomendable solamente si las variables de entrada fueron estandarizadas antes de ser utilizadas (Johnson y Wichern, 2007).

### 1.1.4 Limitaciones del análisis discriminante lineal

Una de las limitaciones más consideradas que presenta la técnica de discriminante lineal, ocurre cuando se presentan datos con ausencia de normalidad o que incumplen el supuesto de homocedasticidad, por lo que al ser una técnica pensada para separar linealmente los datos puede darse el caso de que el modelo generado presente mayor error en la clasificación y discriminación de los sujetos, puesto que los mismos no pueden ser separados. Véase figuras 1 y 2 (Elhabian y Farag, 2009):

Interfaz de usuario gráfica

Descripción generada automáticamente

**Figura 1. Caso ideal para la separación lineal entre dos clases.**

Referencia: Elaboración propia.

Interfaz de usuario gráfica

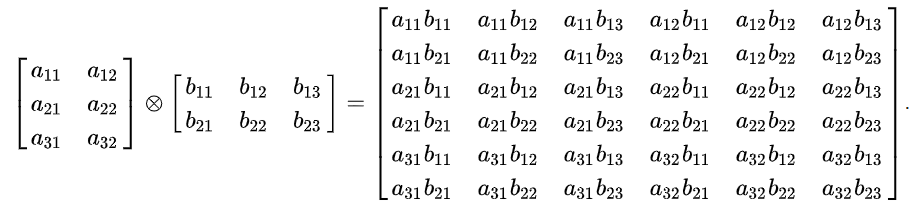
Descripción generada automáticamente

**Figura 2. Casos linealmente no separables.**

Referencia: Adaptado de: *“A Tutorial on Data Reduction Linear Discriminant Analysis (LDA)”* por Elhabian y Farag. 2009. CVIP Lab, Universidad de Louisville.

### 1.2.1 Expansión de Kronecker

El producto de Kronecker es definido como una operación sobrre dos matrices de cualquier tamaño A y B tal que el resultado siempre va a ser una matriz nueva que se define como se muestra a continuación:



## **1.2 Antecedentes**

Yang, Liu, Nie y Liu (2020) Proponen una variante para selección de características y reducción de dimensionalidad basada en LDA y norma L-1,L-2, agregando la implementación de un método de optimización iterativo que recalcula los valores de la separación entre clases para las normas.

Chandrasekar y Geetha (2021) Proponen el uso de un marco de trabajo basado en LDA con diagonal Eigenvalues (LDA-DE) para preservar información en el manejo de grandes volumenes de caracteristicas en imagenes de vehiculos en movimiento. Usaron el F1-Score y Accuracy para evaluar.

Singh, Pal y Dahiya. (2021) Proponen un método para la clasificación de perturbaciones en la calidad de energía eléctrica usando LDA para reducir dimensionalidad y eliminar características innecesarias.

Niu y Ma (2021) Toman el algoritmo de Locally Linear Embedding y hacen el aporte de expandir cada vector de datos usando el producto Kronecker para añadir la componente polinomial y extraer solo las características que mejoran la función discriminante.

Kim, Lee y Liang, Z. (2022) Hacen uso del discriminante lineal con Kernels para ilustrar cómo se comportan los mismos datos con los kernels: Polinomial y Gaussiano con expansión de polinomial Hermite.

Qu y Pei (2024) realizaron una revisión del estado del arte de la técnica, donde se abarcan cerca de 30 variaciones y mejoras que se le han aplicado al discriminante lineal a lo largo de los años, concluyen su revisión haciendo mención sobre expandir la investigación usando métodos Kernels para el problema de la no linealidad.

# 2 Definición del problema

¿Cómo se puede mejorar el discriminante lineal para la separación de los casos en conjuntos de datos no separables linealmente, que pueda ser interpretable y no resulte computacionalmente costoso?

# 3 Propuesta

Implementar una expansión de las características de la matriz de entradas para el discriminante lineal clásico mediante el producto de Kronecker para mejorar la separabilidad de los sujetos.

# 4 Justificación

El Análisis Discriminante Lineal (LDA) es una técnica de clasificación y reducción de dimensionalidad creada por Ronald A. Fisher en 1936 con el propósito de encontrar una combinación lineal que logre maximizar la separación entre clases. Pese a los años transcurridos desde su primera aparición, LDA se mantiene como un método ampliamente utilizado en diversos campos como el reconocimiento de patrones, diagnósticos médicos y procesamiento de imágenes debido a su interpretabilidad y eficiencia computacional.

Sin embargo, una de las principales limitaciones del LDA es el hecho de estar ligada al cumplimiento del supuesto de que los datos son separables linealmente, junto con el cumplimiento de la normalidad y la homogeneidad multivariante entre grupos, esto restringe su desempeño en conjuntos de datos donde las clases presentan una estructura más compleja.

Se a exploraron diversos enfoques para abordar las problemáticas mencionadas previamente, entre ellos se incluye el LDA con Kernels (KDA) y el LDA cuadrático (QDA) que expanden el espacio de características para mejorar la separabilidad de clases.

No obstante, estos métodos suelen introducir hiperparámetros adicionales que deben ser ajustados para lograr su mejor desempeño respecto de la versión básica, lo que puede incrementar la complejidad del modelo y afectar su interpretabilidad.

En este contexto, la presente investigación propone realizar una expansión de la matriz de caractesisticas basada en el producto de Kronecker con el objetivo de capturar interacciones de mayor orden entre las variables sin recurrir a transformaciones arbitrarias o dependientes de hiperparámetros. Este enfoque ha demostrado ser promisorio en otros modelos de clasificación y aprendizaje automático, pero aún no ha sido explorado en profundidad para el caso del LDA. La integración del producto de Kronecker en el LDA podría mejorar su capacidad de clasificar sin comprometerse en la optimización de los hiperpárametros, permitiendo una clasificación más precisa en problemas donde la separación de clases no es estrictamente lineal.

Por lo tanto, este trabajo no solo busca extender las capacidades del LDA mediante una formulación algebraicamente fundamentada, sino que también contribuye a la exploración de técnicas novedosas que optimicen su rendimiento sin comprometer su interpretabilidad ni eficiencia computacional.

# 5 Hipótesis

El rendimiento del discriminante lineal mejora con la matriz de características aumentada con la expansión de Kronecker.

# 6 Objetivos

## 6.1 Objetivo principal

* Diseñar e implementar un modelo de Discriminante Lineal basado en la expansión de Kronecker

## 6.2 Objetivos específicos

* Evaluar la mejora de la separabilidad de clases con la expansión de Kronecker en LDA
* Comparar la Expansión de Kronecker con otras técnicas de mejora en LDA

# 7 Metodología

Se define un conjunto de bases de datos a modo de benchmark.

Se toma la matriz de características y se expande como lo proponen en el articulo de Niu y Ma (2021) mediante el producto de Kronecker.

Hacer selección de las nuevas características para incorporarlas en el discriminante lineal.

En caso de no poderse con el discriminante clásico, probar con el discriminante lineal genético propuesto por García-Medina, et al. (S.F.).

Evaluar el desempeño del algoritmo en tiempo de procesamiento, precisión y F1-score.

Referencias

Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4, No. 4, p. 738). New York: springer.

Briones-Garduño, J. C. y Ponce, M. (2013). *Mortalidad Materna*. Editorial Alfil, S. A. de C. V.

Chandrasekar, K.S., Geetha, P. A new formation of supervised dimensionality reduction method for moving vehicle classification. *Neural Comput & Applic* **33**, 7839–7850 (2021). <https://doi.org/10.1007/s00521-020-05524-z>

Elhabian,S. y Farag, A. A.(2009). *A Tutorial on Data Reduction Linear Discriminant Analysis (LDA)* [Diapositivas de PowerPoint]. CVIP Lab, Universidad de Louisville.

García-Medina, V.-D., Acosta-Mesa, H.-G., López-Lobato, A.-L., & Mezura-Montes, E. (s.f.). Genetic Lineal Discriminant Analysis GLDA. [Manuscrito no publicado]. Artificial Intelligence Research Institute, University of Veracruz.

Johnson, R. A., y Wichern, D. W. (2007). Applied multivariate statistical analysis (6ta ed.). Pearson.

Kim, J., Lee, Y., & Liang, Z. (2022). The geometry of nonlinear embeddings in kernel discriminant analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *45*(4), 5203-5217.

Li, Y., Zhou, R. G., Xu, R., Hu, W., & Fan, P. (2020). Quantum algorithm for the nonlinear dimensionality reduction with arbitrary kernel. *Quantum Science and Technology*, *6*(1), 014001.

Lin, C.-C., & Chen, A.-P. (2004). Fuzzy discriminant analysis with outlier detection by genetic algorithm. *Computers & Operations Research*, *31*(6), 877–888. <https://doi.org/10.1016/s0305-0548(03)00040-6>

McLachlan, G. J. (2004). Discriminant Analysis and Statistical Pattern Recognition. John Wiley & Sons.

Niu, G., & Ma, Z. (2021). Local quasi-linear embedding based on kronecker product expansion of vectors. Journal of Intelligent & Fuzzy Systems, 41(1), 2195-2205.

Silalahi, D. D., Reaño, C. E., Lansigan, F. P., Panopio, R. G., y Bantayan, N. C. (2016). Linear Discriminant Analysis vs. Genetic Algorithm Neural Network with Principal Component Analysis for Hyperdimensional Data Analysis: A study on Ripeness Grading of Oil Palm (Elaeis guineensis Jacq.) Fresh Fruit. *Philippine Statistician*, 65(2).

Singh, G., Pal, Y., & Dahiya, A. K. (2023). Classification of power quality disturbances using linear discriminant analysis. Applied Soft Computing, 138, 110181.

Qu, L., & Pei, Y. (2024). A Comprehensive Review on Discriminant Analysis for Addressing Challenges of Class-Level Limitations, Small Sample Size, and Robustness. Processes, 12(7), 1382.

Xu, Y., & Li, E. (2024). Robust locally nonlinear embedding (RLNE) for dimensionality reduction of high-dimensional data with noise. Neurocomputing, 127900..

“Lis de Veracruz: Arte, Ciencia, Luz”

**www.uv.mx**