Linear Discriminant Analysis - LDA

1. **Giới thiệu**

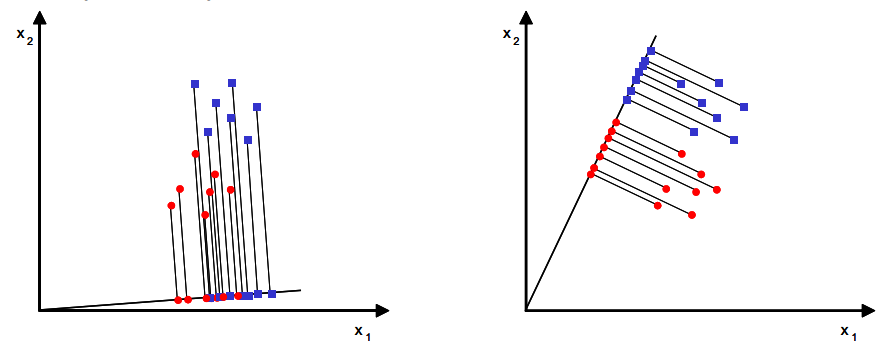
LDA (Linear Discriminant Analysis hay còn gọi là Fisher’s linear discriminant) là phương pháp phân tích trong thống kê được dùng rất nhiều trong data mining để phân loại các đối tượng (object) vào các nhóm dựa trên việc đo lường các đặc trưng của đối tượng; bằng một phép chiếu tuyến tính từ không gian nhiều chiều vào một không gian ít chiều hơn nhưng vẫn giữ được các chất lượng của mô hình. LDA được sử dụng thành công để giảm chiều dữ liệu cho nhiều lĩnh vực chẳng hạn như nhận dạng giọng nói,nhận diện khuôn mặt, nhận dạng đối tượng phức tạp, hay đa phương tiện trong phục hồi hình ảnh...

1. **Mục tiêu của phương pháp LDA**

Mục tiêu của LDA là tìm tập các thuộc tính tốt nhất để mô tả đối tượng. Vấn đề này còn gọi là trích chọn đặc trưng (Feature Extraction) hay lựa chọn đặc trưng (Feature Selection). Mục đích là từ tập các thuộc tính mô tả đối tượng ban đầu (thường rất lớn, có thể hàng trăm thuộc tính) tìm ra tập các thuộc tính tốt nhất để biểu diễn cho đối tượng nhằm làm giảm số chiều biểu diễn đối tượng từ đó làm giảm chi phí tính toán của các mô hình khai phá dữ liệu nhưng chất lượng của mô hình không bị giảm đi. Việc làm này có vẻ tương đồng với PCA, nhưng LDA đã thu gọn tập dữ liệu với các lớp thông tin càng tách biệt nhau càng tốt, để từ đó suy ra một mô hình tốt hơn để phân lớp các đối tượng.

**3. Phương thức hoạt động của LDA**

Trước hết, ta tìm hiểu về việc phân tách nhị phân dùng LDA. Giả sử ta có một tập gồm N mẫu học bao gồm các vector cột d chiều {xi1, xi2…,xid} với i =1…n, trong đó n1 mẫu trong số đó thuộc lớp ω1 và nằm trong tập con x1, và n2 mẫu trong đó thuộc lớp ω2 và nằm trong tập con x2. Nhiệm vụ của LDA là tìm một phép chiếu w sao cho tập mẫu x được chiếu lên một đường thẳng một cách phân biệt nhất.



Ta thấy phép chiếu bên trái đã có sự tối đa hóa về việc phân chia dữ liệu độc lập nhau. Để có thể tìm được một vector chiếu tốt nhất, ta cần xác định một chuẩn mực để phân chia dữ liệu.

Giả sử ta có vector d chiều w, tích vô hướng y = wTx biểu diễn hình chiếu của vector x lên phương w. Ta sẽ tìm một phương chiếu w nhằm tối ưu hóa độ phân biệt mẫu giữa các mẫu thuộc 2 lớp C1 và C2. Điều này tương đương với việc giảm số chiều của vector đặc trưng xuống 1 chiều.

Gọi μilà vector trung bình của cácmẫu tương ứng với lớp C1 và C2 và tương ứng là trung bình mẫu được chiếu lên phương w

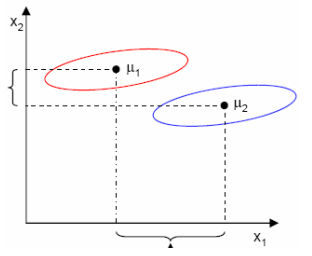


Trong đó y là hình chiếu của x lên w, yi là tập các hình chiếu của các xi lên w

Và hàm mục tiêu có thể được lựa chọn như sau:



Đây có thể xem là một độ đo cho tính phân biệt giữa hai tập y1 và y2. Tuy nhiên, để thu được sự phân biệt tốt giữa hai tập dữ liệu khi chiếu lên w, ta cần giá trị sai khác J(w) này rất lớn so với độ lệch chuẩn nội tại của mỗi tập (hay nói cách khác là độ rộng các đám mây của mẫu). Quan sát ví dụ sau để có thể thấy được kết quả phân lớp không tối ưu khi dùng hàm mục tiêu J(w) như trên.

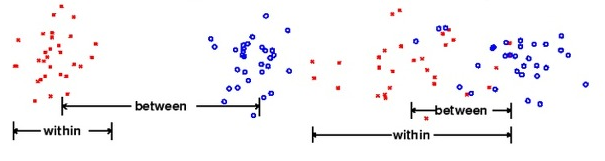


Theo phương x1, rõ ràng khoảng cách đạt cực đại nhưng phép chiếu lại có sự chồng chập dữ liệu và phương x2 cho kết quả chiếu dữ liệu độc lập hơn.

Giải pháp cho vấn đề này được đề xuất bởi Fisher, với định nghĩa hàm mục tiêu mới: độ tán xạ (scatter). Để phân tách lớp dữ liệu, người ta dùng một tỉ lệ tán xạ nội lớp và ngoại lớp

**

Minh họa bởi hình vẽ sau



Phân lớp dữ liệu tốt Phân lớp dữ liệu kém

Ta sử dụng định hướng trên để giải quyết bài toán.

Độ biến thiên của lớp ω1 khi chiếu lên không gian y:



****là độ biến thiên của hai lớp khi chiếu lên không gian y, hay còn gọi là tán xạ nội lớp (within-class scatter)

Phương pháp LDA sẽ tìm giá trị w để cực đại hóa hàm mục tiêu sau đây:



Để tìm phép chiếu w\* tối ưu, cần phân tích J(w) theo biến w, và người ta đưa ra một số định nghĩa mới như sau.

* Ma trận tán xạ nội lớp (within-class scatter matrix) Sw: là thước đo sự phân tán của các mẫu trong không gian đặc trưng đa biến, và Si là ma trận hiệp phương sai của lớp wi



Vậy ta có:



Suy ra: ****

Với  là ma trận tán xạ nội lớp khi chiếu lên không gian y

* Tương tự:



Trong đó  được gọi là ma trận tán xạ liên hợp (between-class scatter) và  là ma trận tán xạ liên hợp trên không gian y.

Như vậy, hàm mục tiêu được viết lại như sau:



Để tìm giá trị lớn nhất của hàm mục tiêu, ta thực hiện đạo hàm J(w) và cho kết quả bằng 0



Như vậy ta đưa bài toán về tìm trị riêng λ của ma trận: . Giải bài toán tìm trị riêng, kết quả thu được là . Đây là kết quả tìm được của phương pháp LDA đối với trường hợp phân chia 2 lớp.

Sau khi đã tìm được w, mỗi vector x cần nhận dạng sẽ được xử lý như sau: lấy x trừ đi trung bình của mẫu học rồi chiếu lên phương w ta được một giá trị vô hướng, tính khoảng cách từ giá trị vô hướng này với  của mỗi lớp.



Đây chính là khoảng cách từ x đến các cụm tương ứng, và đối tượng x sẽ được gán đến lớp có khoảng cách d bé nhất.

**4. Ví dụ về phân lớp LDA**

Thực hiện tìm phép chiếu LDA cho 2 tập mẫu dữ liếu sau:

* Các mẫu thuộc lớp ω1: X1=(x1,x2)={(4,2),(2,4),(2,3),(3,6),(4,4)}
* Các mẫu thuộc lớp ω2: X2=(x1,x2)={(9,10),(6,8),(9,5),(8,7),(10,8)}

Giá trị trung bình của mỗi lớp :



Ma trận hiệp phương sai của mỗi lớp:



Ma trận tán xạ nội lớp:



Ma trận tán xạ liên hợp:



Thực hiện bài toán tìm trị riêng theo phương pháp LDA:



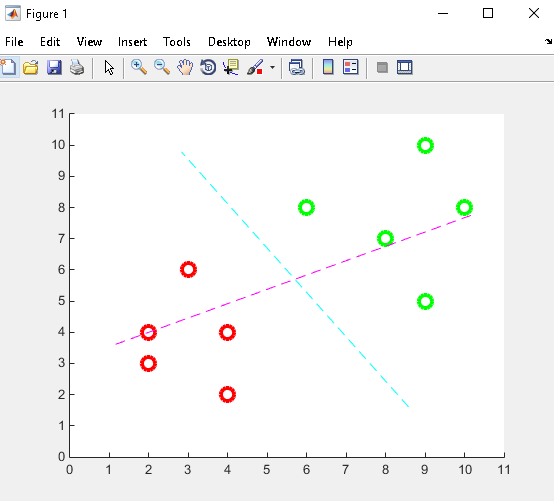
Khi đó:



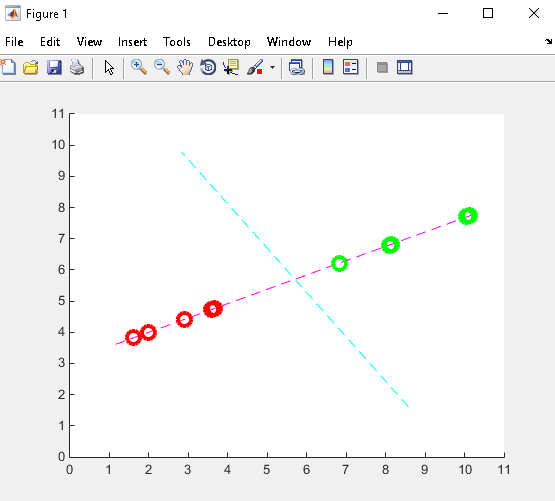
Phương chiếu được lựa chọn ứng với trị riêng λ lớn nhất (trong ví dụ này là λ­2), hoặc có thể tính toán trực tiếp theo công thức:



Các tập mẫu và phương chiếu được biểu diễn trên hình vẽ sau đây:



Trong đó, phương chiếu w1 ứng với đường kẻ màu xanh và phương chiếu w2 ứng với đường kẻ màu hồng. Thực hiện phép chiếu ta thu được kết quả sau:



**5. LDA đa lớp**

Thuật toán LDA được phát triển để giải bài toán phân lớp với số lượng lớp C>2. Ta cần tìm bộ phương chiếu w­i sao cho hình chiếu y của tập dữ liệu x ban đầu phân biệt rõ ràng nhất.



Với:



Giả sử ta có n đặc trưng:



Ma trận tán xạ nội lớp:



Với Ni là số mẫu trong lớp ωi

Ma trận tán xạ liên hợp:



Với N là tổng số mẫu trong tập dữ liệu

Thực hiện hoàn toàn tương tự để thu được các hình chiếu lên phương w:



Bài toán LDA cho C lớp cũng quy về giải tìm trị riêng của ma trận:



Kết quả cuối cùng là một ma trận phương chiếu tối ưu 