

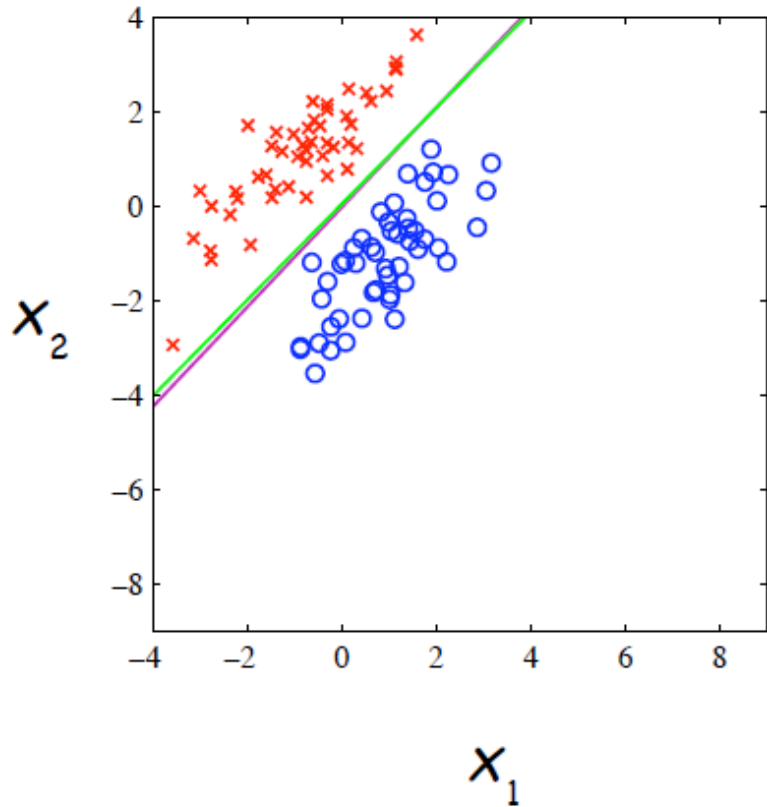
Classification

مدلهای خطی

گروه دایچه . dayche.com



روش‌های تحلیل تابع جداساز خطی



- تابع جداساز خطی

- داده‌های مربوط به دو کلاس توسط یک سطح تصمیم خطی جدا می‌شوند

$$\Rightarrow y(x) = w_1x_1 + w_2x_2 + w_0 \rightarrow \begin{cases} y(x) > 0 \rightarrow C_1 \\ y(x) < 0 \rightarrow C_2 \end{cases}$$

$$y(x) = 0$$

Decision surface

- برای هر نمونه عمل مقایسه کردن را به صورت زیر تعریف می‌کنیم:


$$y(x) = \text{sign}(w_1x_1 + w_2x_2 + w_0)$$

Nonlinear function

تولید محتوا: وحید محمدزاده ایوقی

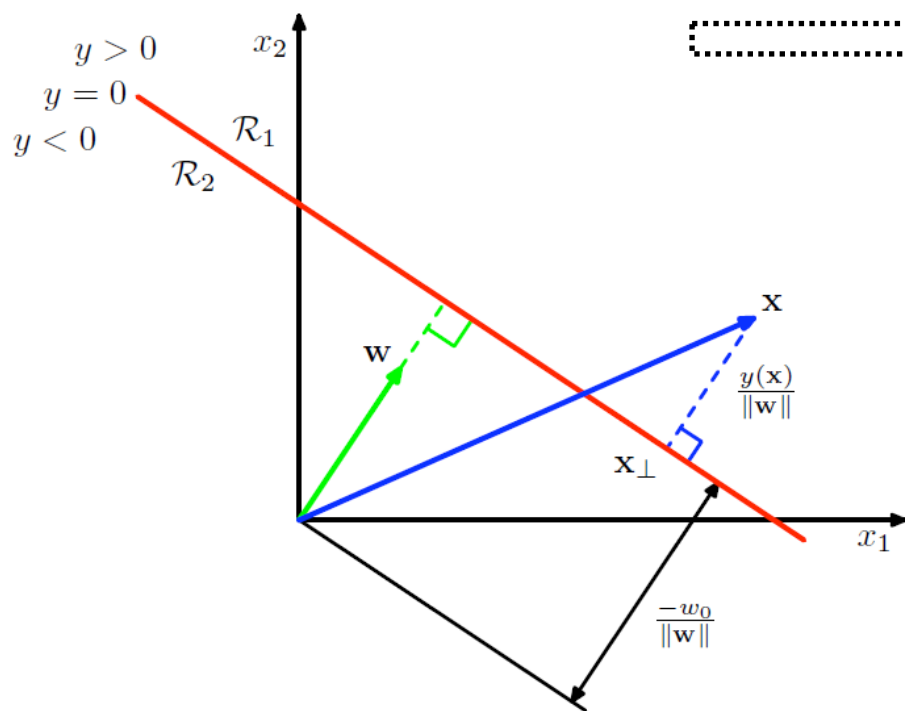
daychegroup 

daychegroup 

dayche.com | گروه دایچه 

روش‌های تحلیل تابع جداساز خطی

• تابع جداساز خطی

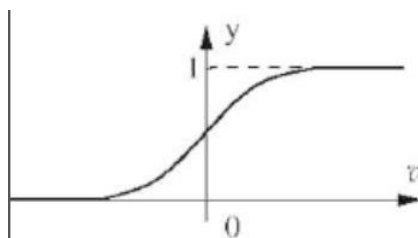


$$y(x) = \text{sign}(W^T x + w_0)$$

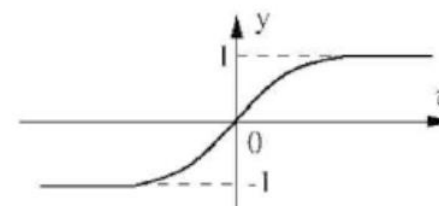
هر مقدار خروجی، یک مقدار گسسته و یا یک بردار احتمال را نشان می‌دهد

$$y(x) = f(W^T x + w_0)$$

Generalized linear model



Log-sigmoid function



Tan-sigmoid function

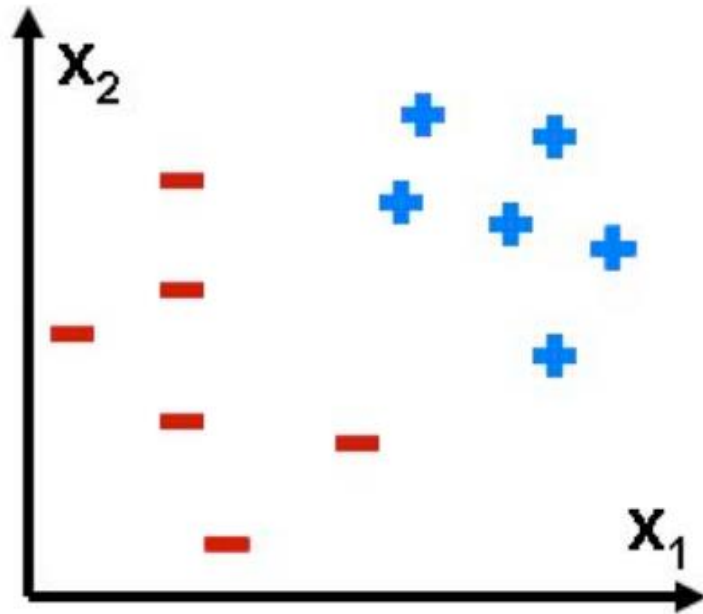
تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

dayche.com | گروه دایچه

آموزش پارامترهای مدل




- الگوریتم پرسپترون
- فرضیات مسئله
- داده های تفکیک پذیر خطی هستند.
- تابع هزینه
- آیا خطای دسته بندی تابع هزینه مناسبی است؟ (چرا؟)

$$y(x) = W^T x + w_0 \quad \Leftrightarrow \quad \begin{cases} y(x_n) \geq 0 \rightarrow t_n = 1 \\ y(x_n) < 0 \rightarrow t_n = -1 \end{cases} \quad \Leftrightarrow \quad t_n y(x_n) \geq 0$$

تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایچه 



- آموزش پارامترهای مدل:

$$W^* = \arg \max \sum_{n=1}^N t_n y(x_n) \rightarrow W^* = - \arg \min \sum_{x_n \in M} t_n y(x_n), \quad y(x_n) = W^T x_n$$

- گرادیان نزولی:

- برای داده‌های کلاس 2، بردار وزن به گونه‌ای تغییر می‌کند تا تصویر این نقاط بر روی سطح تصمیم منفی‌تر باشد
- برای داده‌های کلاس 1، بردار وزن به گونه‌ای تغییر می‌کند تا تصویر این نقاط بر روی سطح تصمیم مثبت‌تر باشد

$$W^+ = W^- + \eta \sum_{x_n \in M} x_n t_n$$

تولید محتوا: وحید محمدزاده ایوقی

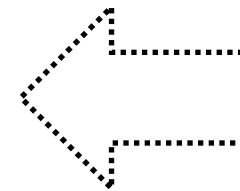
daychegroup

daychegroup

dayche.com | گروه دایکه

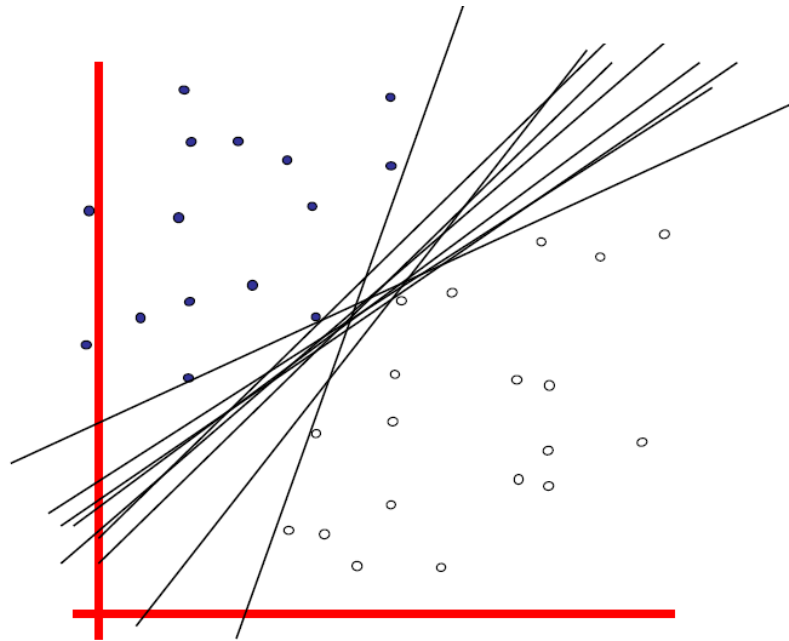


- نرخ یادگیری باید از ثابت لیپشتیز تابع هزینه کوچک تر باشد تا همگرایی تضمین شود.
- در صورتی که داده ها تفکیک پذیر خطی باشند و نرخ یادگیری مناسب انتخاب شده باشد، الگوریتم قطعا همگرا خواهد شد



- مقدار اولیه وزن‌ها
- ترتیب داده‌ها
- نرخ یادگیری

- مقدار اولیه وزن‌ها
- ترتیب داده‌ها
- نرخ یادگیری




- مشکل: با تغییر وزندهی اولیه، سطوح تصمیم‌گیری متفاوتی حاصل می‌شود!
- سطح بهینه تصمیم‌گیری، سطحی است که خطای تعمیم دهی کمتری داشته باشد

تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایچه 




- تفکیک‌پذیری خطی
 - آیا تفاوت بین تفکیک‌ناپذیر بودن خطی و سرعت پایین همگرایی قابل بررسی است؟
 - اگر نرخ یادگیری مناسب باشد، یک روش عملیاتی بررسی تفکیک‌پذیر خطی بودن دو کلاس استفاده از الگوریتم پرسپترون است. (چگونه؟)
 - به لحاظ تئوری، معیار FDR دو کلاس باید بزرگ باشد.
 - Convex hull متناظر با دو کلاس دارای تلاقی نباشند.
- دو کلاس جدایی‌ناپذیر خطی
 - Pocket algorithm

تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

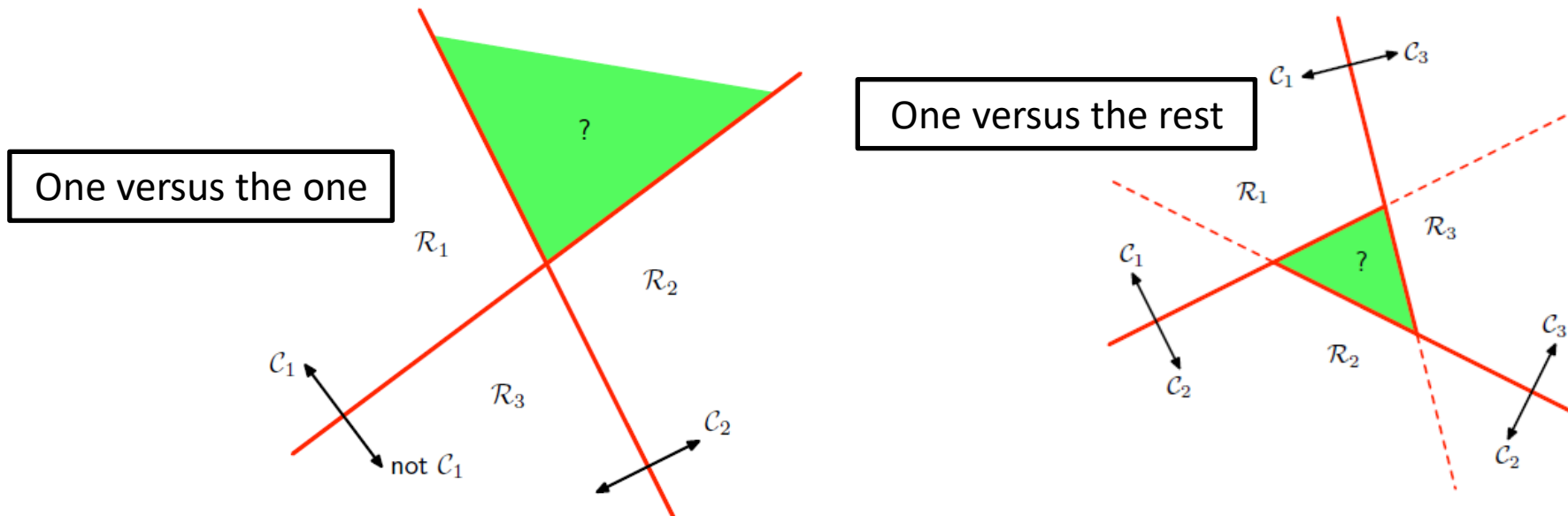
dayche.com | گروه دایکه 

مسئله چند کلاسه

• سطح تصمیم

- اگر نگاشت هر نقطه بر روی سطح تصمیم با 0 مقایسه شود، آیا می‌توان سطح تصمیم بهینه‌ای بدست آورد؟ (خیر! چرا؟)
- دو رویکرد کلی برای تفکیک چند کلاس وجود دارد (تفکیک یک کلاس نسبت به بقیه کلاس‌ها، تفکیک یک کلاس نسبت به تک‌تک


کلاس‌ها)



تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

مسئله چند کلاسه

• راه حل

• برای هر کلاس یک تابع جداساز باید تعریف شود

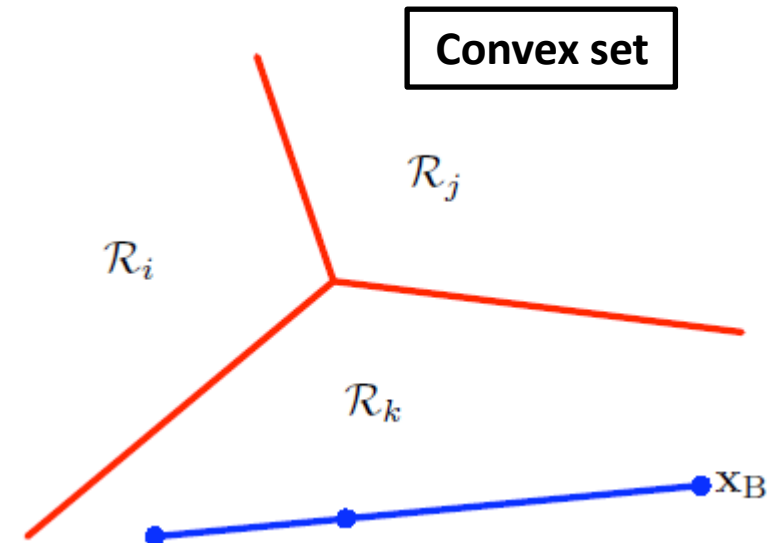
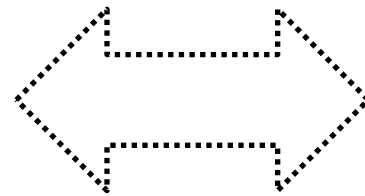
$$y_k(x) = W_k^T x, \quad k = 1, 2, \dots, M$$

$$y_i(x) > y_j(x)$$

کلاس i

$$y_i(x) = y_j(x)$$

Decision surface



تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

dayche.com | گروه دایچه

الگوریتم پرسپترون - چند کلاسه

- تعمیم الگوریتم پرسپترون به حالت چند کلاسه مشکل است.

- Kesler's construction

- به صورت همزمان وزن‌های تابع جداساز باید آموزش ببینند؟

$$C_i = [0 \quad 0 \quad \dots \quad 1 \quad 0]$$

- تحت چه شرایطی مسئله همگرا خواهد شد؟


المان i -ام

کلیه الگوریتم‌های بیان شده بر روی نگاشتی از فضای ورودی نیز قابل اعمال است

تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایکه 

الگوریتم حداقل مربعات



- نیازی به فرضیات مطرح شده برای الگوریتم پرسپترون نیست.
- بردار خروجی در این حالت به صورت یک کد باینری در نظر گرفته می‌شود.

$$t_n = [0 \quad 0 \quad \dots \quad 1 \quad 0]$$

$$C^* = \arg \max y_i \quad T = [t_1 \quad t_2 \quad \dots \quad t_{n-1} \quad t_n]$$

$$y_1(x) = W_1^T x$$

.

.

.

$$y_n(x) = W_n^T x$$

$$Y_{n \times k} = X_{n \times d} \hat{W}_{d \times k} \rightarrow E_D(W) = \frac{1}{2} \text{trace}((XW - T)^T (XW - T))$$

MSE

$$W^* \arg \min(E_D(W)) \iff W^* = (X^T X)^{-1} X^T T$$

تولید محتوا: وحید محمدزاده ایوقی

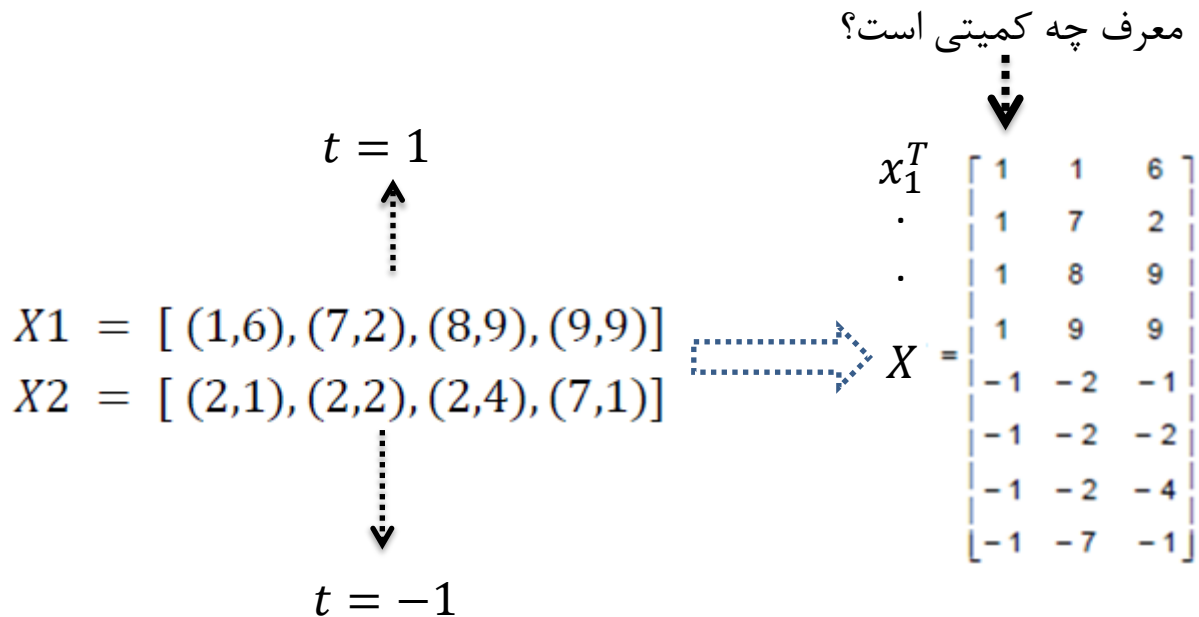
daychegroup

daychegroup

dayche.com | گروه دایچه

دسته‌بندی - مثال

• مثال - پرسپترون و حداقل مربعات



$$\eta = 0.1, W_0 = [0.1, 0.1, 0.1]^T$$

$$y_1 = W_0^T x_1 = 0.8 > 0 \quad \text{No update}$$

$$y_5 = W_0^T x_5 = -0.4 < 0 \rightarrow W_1 = W_0 + 0.1 \begin{bmatrix} -1 \\ -2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ -0.1 \\ 0 \end{bmatrix}$$

No update


$$W^* = [-3.5, 0.7, 0.3]^T$$

بعد از تکرار ۱۷۶ گام برای الگوریتم

تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

مثال - پرسپترون و حداقل مربعات

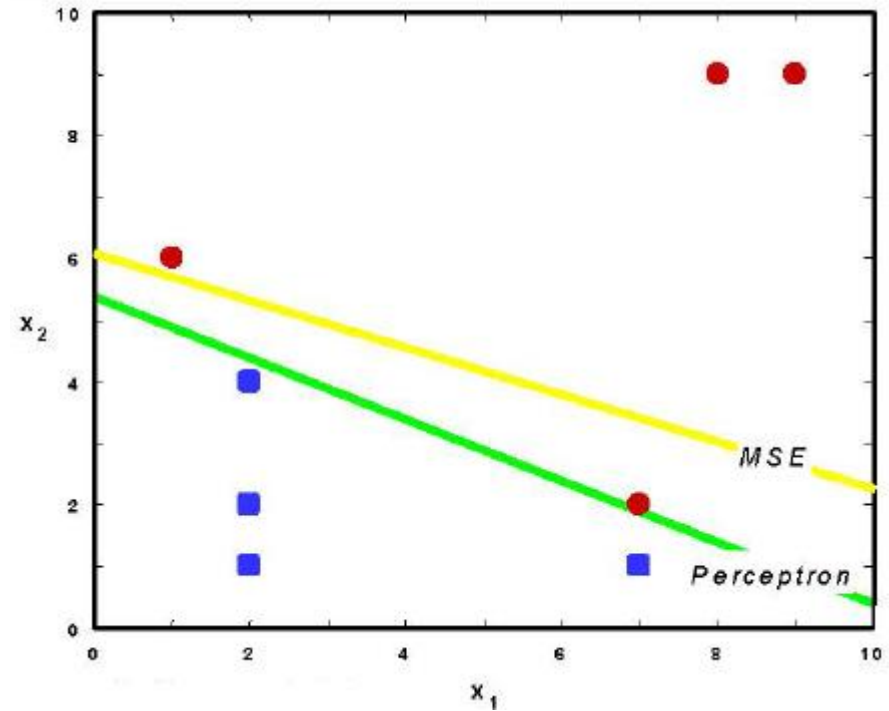
$$X1 = [(1,6), (7,2), (8,9), (9,9)]$$

$$X2 = [(2,1), (2,2), (2,4), (7,1)]$$



$$= \begin{bmatrix} 1 & 1 & 6 \\ 1 & 7 & 2 \\ 1 & 8 & 9 \\ 1 & 9 & 9 \\ -1 & -2 & -1 \\ -1 & -2 & -2 \\ -1 & -2 & -4 \\ -1 & -7 & -1 \end{bmatrix}$$


$$T = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \rightarrow W^* = (X^T X)^{-1} X^T T = \begin{bmatrix} -1.2 \\ 0.07 \\ 0.2 \end{bmatrix}$$



تولید محتوا: وحید محمدزاده ایوقی

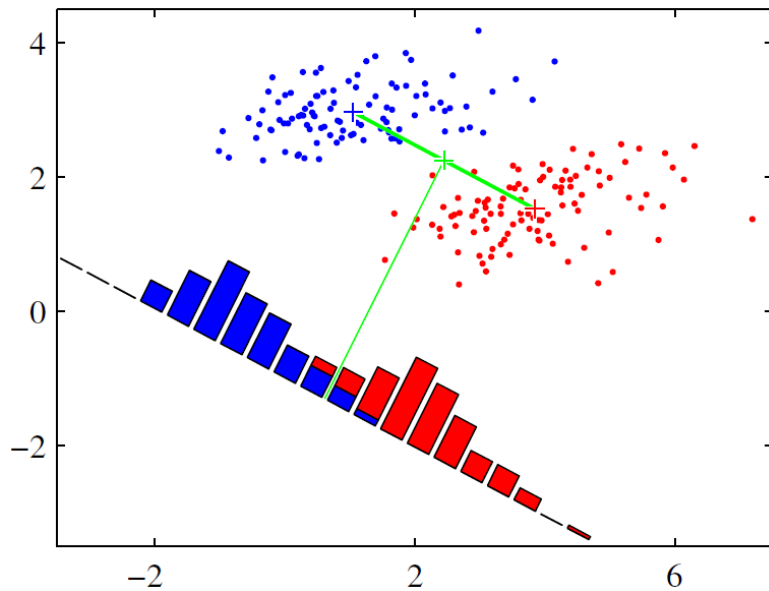
daychegroup 

daychegroup 

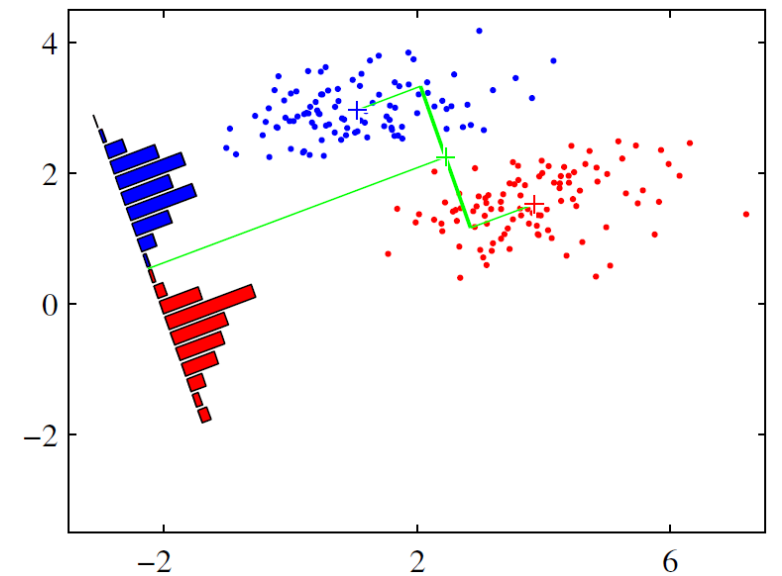
dayche.com | گروه دایکه 



- همانطور که گفته شد، دسته بندی بر اساس تصویر داده بر روی سطح تصمیم انجام می‌پذیرد.



راه حل:
بردار وزن‌ها طوری محاسبه شوند
که در فضای جدید، میزان
تفکیک‌پذیری دو کلاس بالا باشد.



تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

dayche.com | گروه دایکه

$$y(x) = W^T x \rightarrow \begin{cases} m_1 = W^T m_1 \\ m_2 = W^T m_2 \end{cases}$$

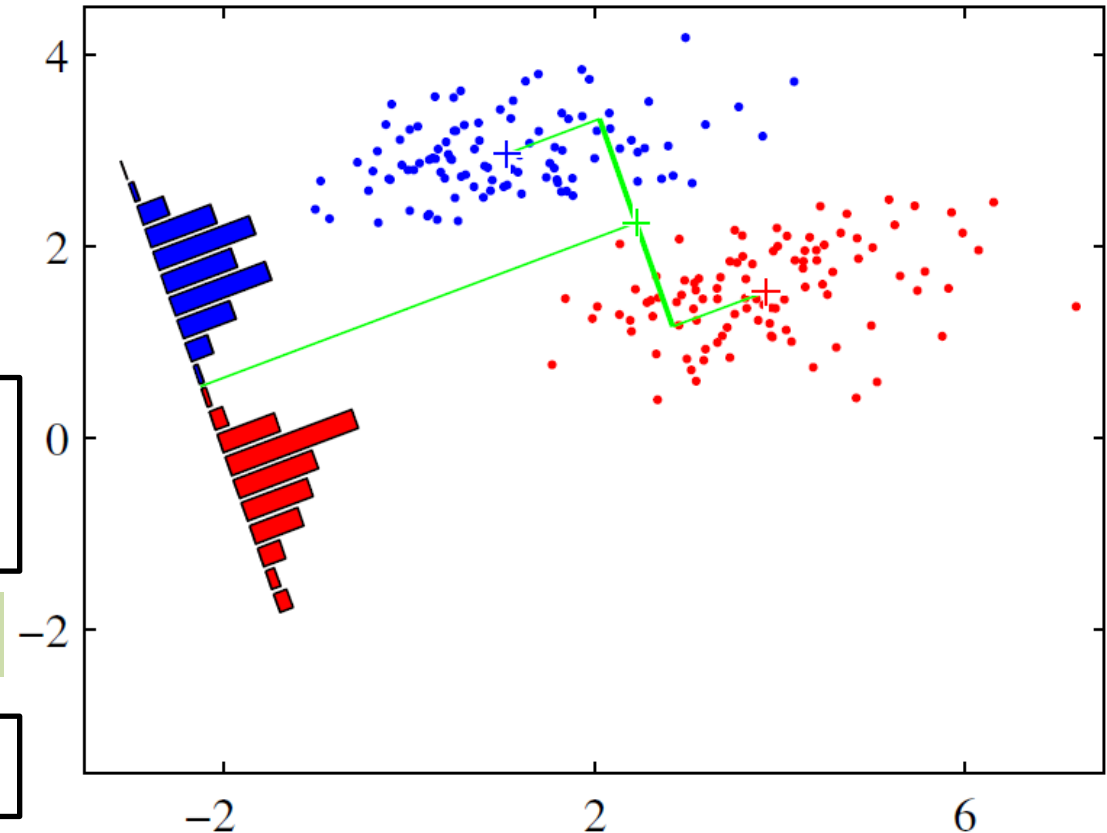
$$m_1 - m_2 = W^T (m_1 - m_2)$$

هدف افزایش میانگین دو کلاس نگاشت داده شده است. اگر بردار وزن‌ها را بسیار بزرگ انتخاب کنیم این مهم تحقق می‌یابد.

آیا این رویکرد درست است؟ چرا؟

$$W^* = \arg \max W^T (m_1 - m_2), s.t. |W| < 1$$


Constrained optimization



تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

- در بسیاری از موارد پراکندگی داده‌ها به صورت شعاعی نیست. در نتیجه معیار معرفی شده قابلیت یافتن وزن‌های مناسب را نخواهد داشت.

ایده Fisher

$$J(W) = \frac{(m_1 - m_2)^2}{S_{W_1}^2 + S_{W_2}^2} \rightarrow W^* = \arg \max J(W)$$

$$S_B = (m_1 - m_2)(m_1 - m_2)^T$$

$$S_W = \sum_{x_k \in C_1} (x_k - m_k)(x_k - m_k) + \sum_{x_k \in C_2} (x_k - m_k)(x_k - m_k)$$

$$J(W) = \frac{W^T S_B W}{W^T S_W W}$$

$$W^* \propto S_W^{-1} (m_1 - m_2)$$

Fisher discriminant analysis (FDA)

آنالیز خطی فیشر – مثال

• آنالیز خطی فیشر

$$C_1 = [(1, 2), (2, 3), (3, 3), (4, 5), (5, 5)]$$

$$C_2 = [(1, 0), (2, 1), (3, 1), (3, 2), (5, 3), (6, 5)]$$

$$\mu_1 = [3, 3.6], \mu_2 = [3.3, 2]$$

حل:


$$S_1 = 4\Sigma_{c_1} = \begin{bmatrix} 10 & 8 \\ 8 & 7.2 \end{bmatrix}, S_2 = 5\Sigma_{c_2} = \begin{bmatrix} 17.3 & 16 \\ 16 & 16 \end{bmatrix} \Rightarrow S_w = \begin{bmatrix} 27.3 & 24 \\ 24 & 23.2 \end{bmatrix}$$

$$W^* = S_w^{-1}(\mu_1 - \mu_2) = \begin{bmatrix} -0.79 \\ 0.89 \end{bmatrix}$$

تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایکه 

آنالیز خطی فیشر - چندکلاسه

➤ تعمیم FDA به مسئله چند کلاسه

$$S_W = \sum_k S_{W_k}, S_{W_k} = \sum_{x_n \in C_k} (x_n - m_k)(x_n - m_k)^T$$

$$S_B = \sum_k (m_k - m)(m_k - m)$$

میانگین کل داده های آموزش

$$J(W) = \text{trace}((W^T S_W W)^{-1} (W^T S_B W))$$

$$W^* \in \text{eigenvectors}(S_W^{-1} S_B)$$

اگر K کلاس موجود باشد، بردار W شامل $k - 1$ بردار ویژه ماتریس $S_W^{-1} S_B$ خواهد بود (چرا؟). بنابراین روش FDA نگاشتی از فضای ورودی را بدست می‌دهد که حداکثر بُعد فضای جدید برابر با $k - 1$ خواهد بود.

تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایکه 