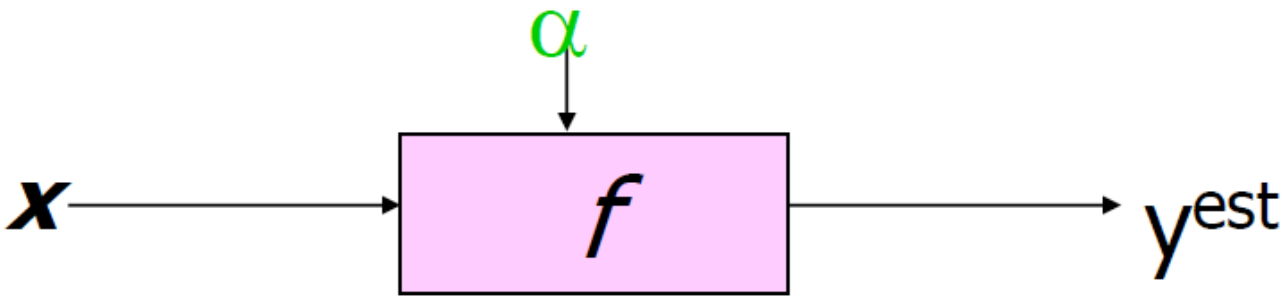
ماشین بردار پشتیبان

dayche.com . گروه دایکه

ماشین بردار پشتیبان

- ماشین بردار پشتیبان
- اولین بار برای دسته‌بندی داده‌ها توسط Vapnik معرفی شد.
- دسته‌بندی کننده خطی

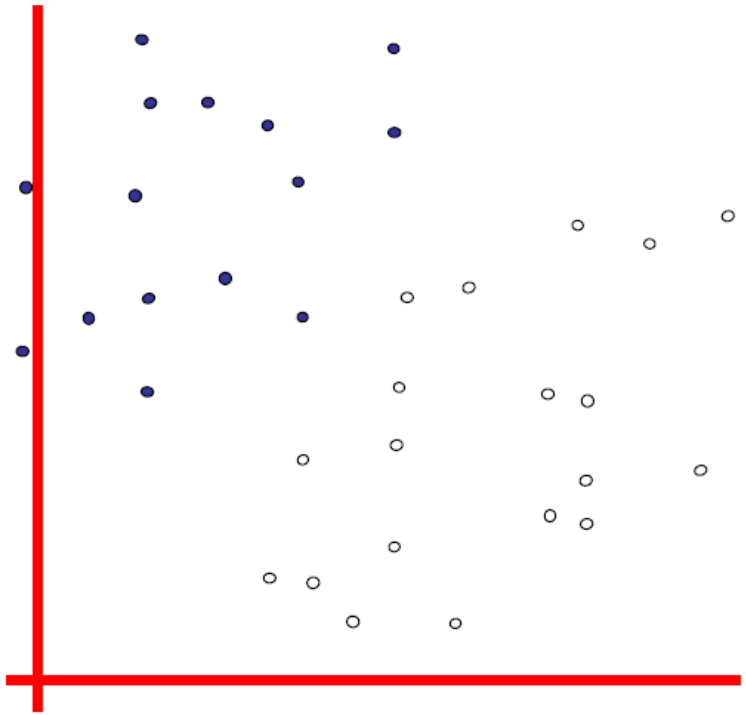


$$y_{est} = f(wx + b)$$

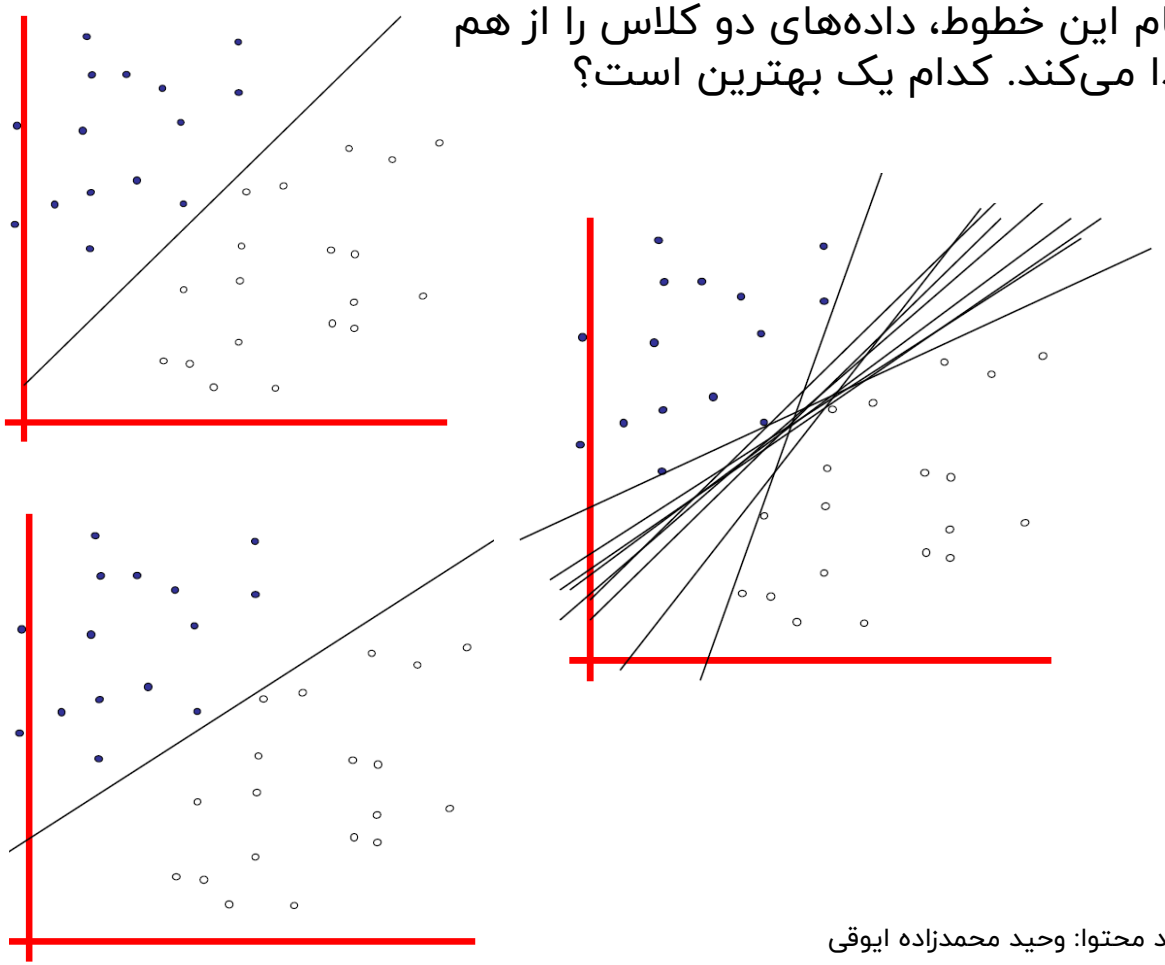
تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

دسته‌بندی کننده خطی

داده‌ها زیر چطور دسته‌بندی می‌شوند؟



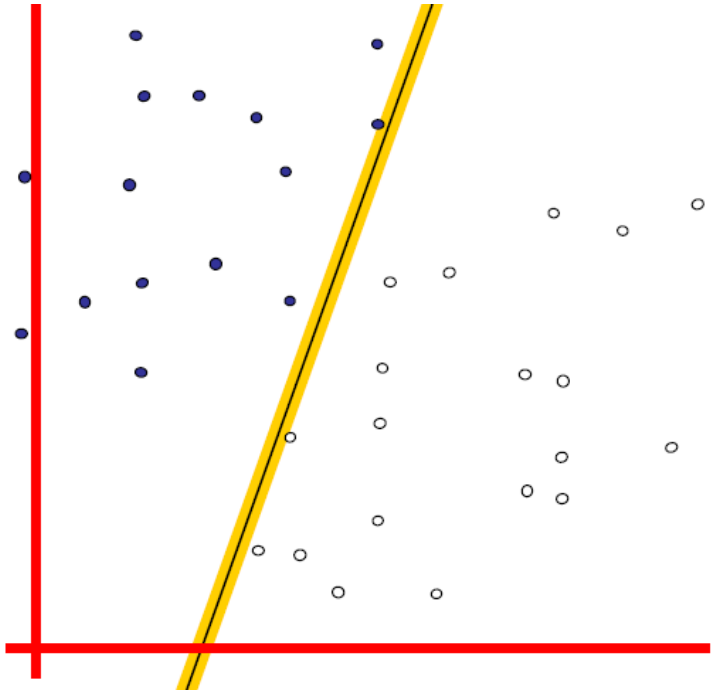
تمام این خطوط، داده‌های دو کلاس را از هم جدا می‌کند. کدام یک بهترین است؟



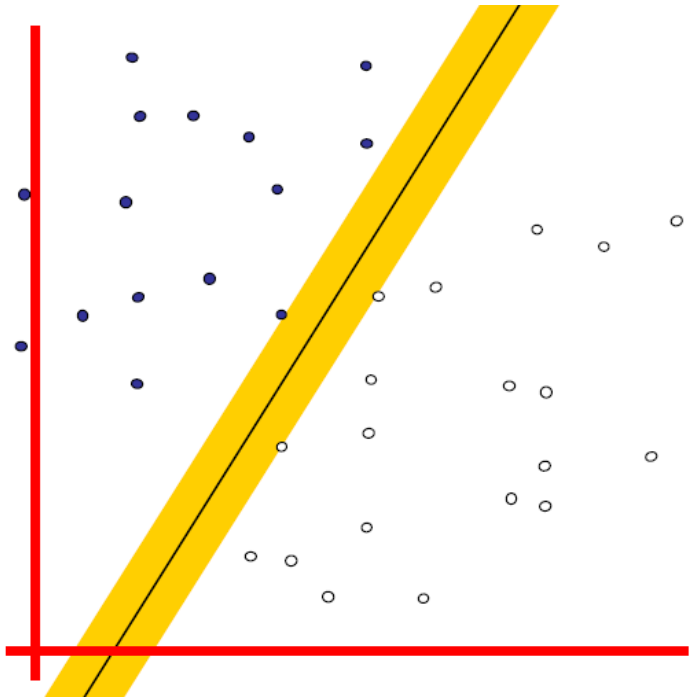
تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

بهترین دسته‌بندی کننده‌ی خطی

دسته‌بندی کننده خطی بیشینه حاشیه – ماشین بردار پشتیبان خطی

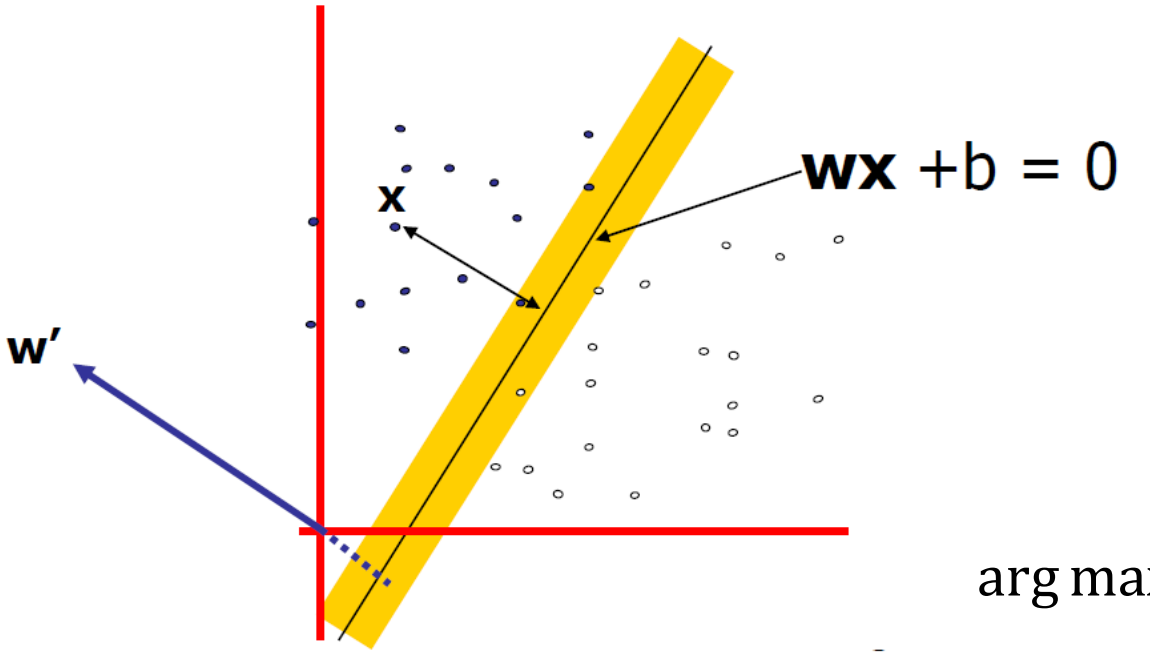


خطی را انتخاب می‌کنیم که کم‌ترین خطای تعمیم‌دهی را داشته باشد
بهترین خط، خطی‌ست که بیشترین فاصله را از مرز داده‌های دو کلاس داشته باشد



تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

ماشین بردار پشتیبان



- ماشین بردار پشتیبان

$$d(x) = \frac{|Wx + b|}{|W|_2}$$

$$W^* = \arg\max \arg\min d(x)$$

$$\arg\max \arg\min d(x)$$

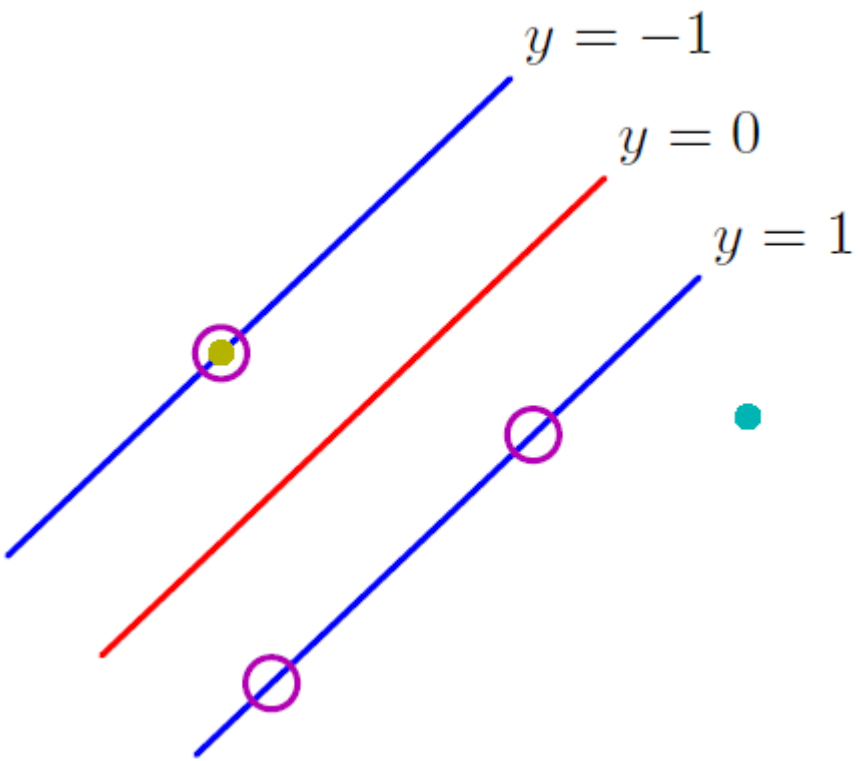
$$y_n(w^T x_n + b) \geq 0$$

$$W^* = \arg\min \frac{1}{2}|W|^2$$

$$y_n(w^T x_n + b) \geq \epsilon$$

تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
گروه دایکه | dayche.com

ماشین بردار پشتیبان



تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

بهینه‌سازی مقید

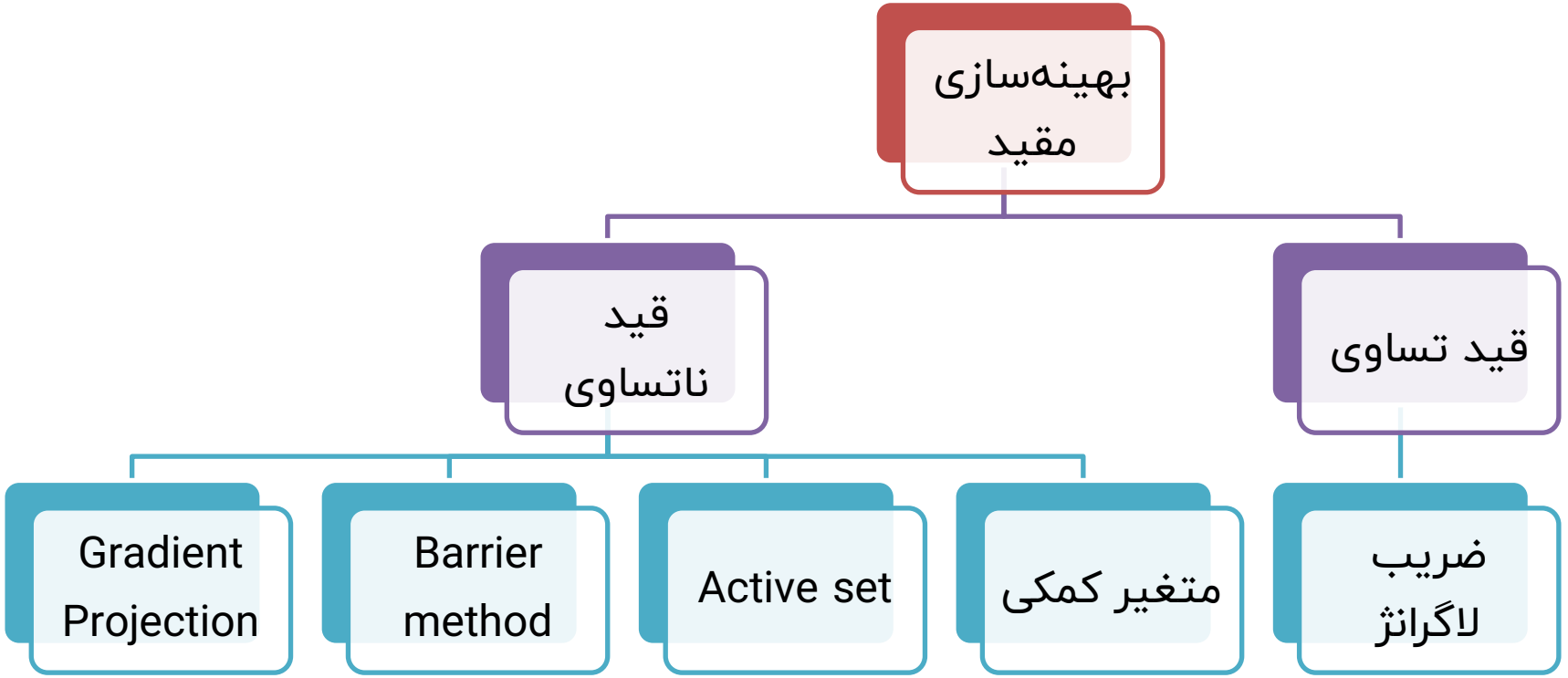
- تعریف
- پارامتر بهینه‌سازی هر مقدار را نمی‌تواند بپذیرد.

$$x^* = \arg\min f(x)$$
$$h_j(x) = 0, j = 1, 2, \dots, k$$
$$g_i(x) < 0, i = 1, 2, \dots, m$$

$$x^* = \arg\min x_1^2 + (x_2 - 1)^2$$
$$-2x_1^2 + x_2 = 4$$
$$x^* = \arg\min x_1^2 + x_2^2$$
$$-x_1 + x_2 < 4$$
$$W^* = \arg\min \frac{1}{2}|W|^2$$
$$y_n(w^T x_n + b) \geq \epsilon$$

تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

بهینه‌سازی مقید

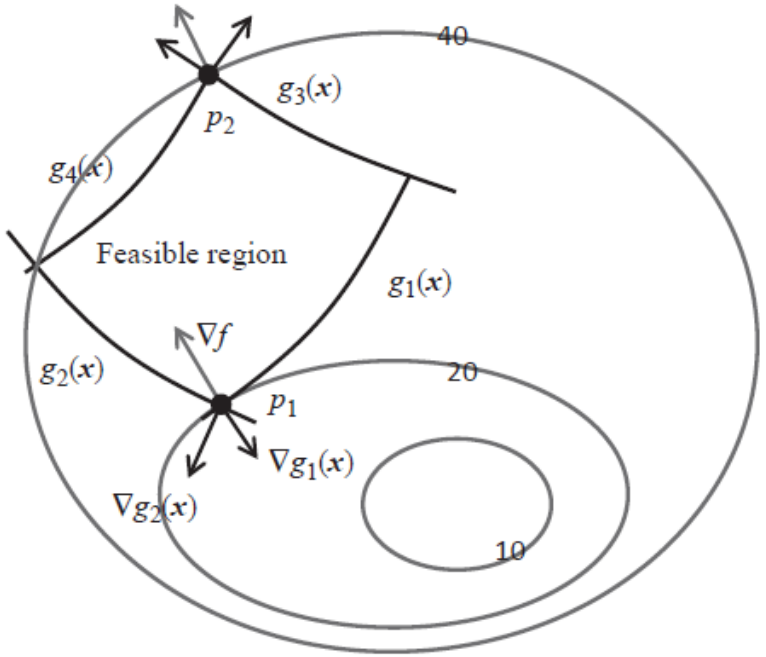


تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

بهینه‌سازی مقید

- روش Active set

در نقطه بهینه، گرادیان تابع می‌تواند بر حسب ترکیب خطی گرادیان هر مجموعه بیان شود.



$$\nabla f = -\sum_{i=1}^{m} \lambda_i \nabla g_i(x)$$

شرایط KKT

$$\lambda_i g_i = 0, \qquad \lambda_i \geq 0$$

- حل مسئله با فرض اکتیو بودن تمام قیود آغاز می‌شود (الزام نیست)
- ضرایب لاگرانژ محاسبه می‌شوند.
- اگر مقادیر ضرایب لاگرانژ مثبت بود، فرض اکتیو بودن مجموعه درست است.

روش Active set – مثال

$$f(x_1,x_2)=x_1^2+x_2^2-14x_1-6x_2$$
$$g_1(x_1,x_2)=x_1+x_2-2\le 0$$
$$g_2(x_1,x_2)=x_1+2x_2-3\le 0$$

$$\nabla f=\begin{pmatrix}2x_1-14\\2x_2-6\end{pmatrix},\nabla g_1=\begin{pmatrix}1\\1\end{pmatrix},\nabla g_2=\begin{pmatrix}1\\2\end{pmatrix}$$

$$\begin{cases}2x_1+\lambda_1+\lambda_2=14\\2x_2+\lambda_1+2\lambda_2=6\end{cases}$$

اگر هیچ قیدی فعال نباشد: $x_1=7,x_2=3\rightarrow\begin{cases}7+3-2=8>0\\7+6-3=10>0\end{cases}$

اگر فقط قید اول فعال باشد: $x_1=\frac{14-\lambda_1}{2},x_2=\frac{6-\lambda_1}{2}\rightarrow\frac{14-\lambda_1}{2}+\frac{6-\lambda_1}{2}-2=0\rightarrow\lambda_1=8,x=\begin{pmatrix}3\\-1\end{pmatrix}$

اگر فقط قید دوم فعال باشد: $x_1=\frac{14-\lambda_2}{2},x_2=3-\lambda_2\rightarrow\frac{14-\lambda_1}{2}+6-\lambda_1-3=0\rightarrow\lambda_4=4,x=\begin{pmatrix}5\\-1\end{pmatrix}$

اگر هر دو قید فعال باشند: $x_1=\frac{14-\lambda_2-\lambda_1}{2},x_2=\frac{6-\lambda_1-2\lambda_2}{2},\lambda_1=20,\lambda_2=-8$

تولید محتوا: وحید محمدزاده ایوقی

daychegroup
daychegroup
dayche.com | گروه دایکه

ماشین بردار پشتیبان خطی

$$W^* = \arg\min \frac{1}{2}|W|^2$$

$$y_n(wx_n + b) \geq 1$$

$$W^* = \arg\min \left(\frac{1}{2}|W|^2 - \sum_{i=1}^{n} \alpha_i (y_i(w^T x_i + b) - 1) \right)$$

$$J = \frac{1}{2} w^T w - \sum_{i=1}^{n} \alpha_i y_i w^T x_i - b \sum_{i=1}^{n} \alpha_i y_i + \sum_{i=1}^{n} \alpha_i$$

$$w = \sum_{i=1}^{n} \alpha_i y_i x_i, \sum_{i=1}^{n} \alpha_i y_i = 0$$

$$b = \frac{1}{N_s} (\Sigma y_n - \sum \alpha_m y_m x_i^T x_j)$$

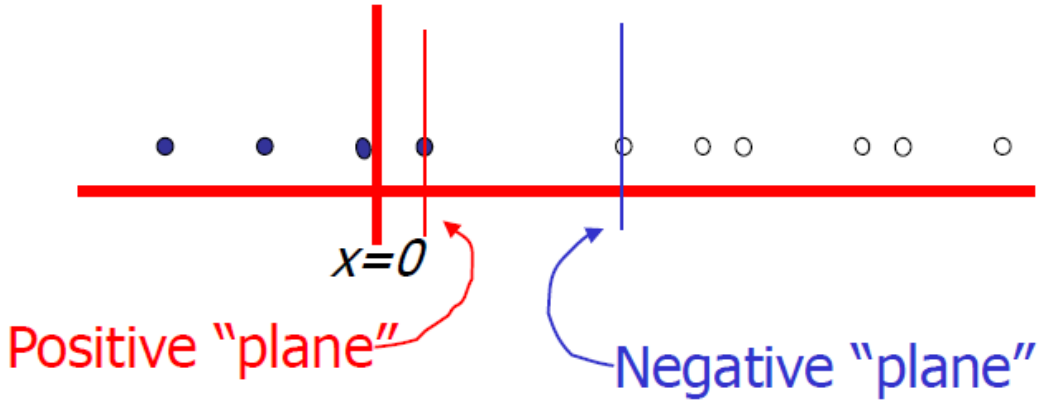
$$J = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j y_i y_j x_i^T x_j$$

$$\sum_{i=1}^{n} \alpha_i y_i = 0, \alpha_i > 0$$

تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

ماشین بردار پشتیبان خطی

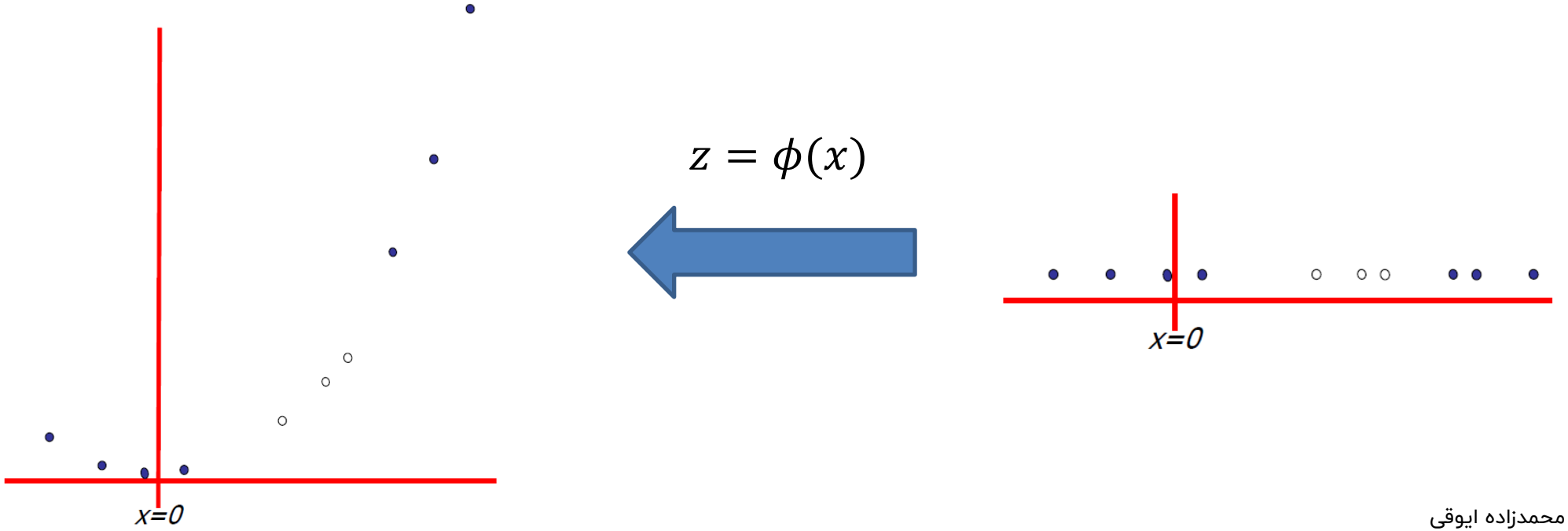
- فرض اصلی تئوری مطرح شده
  - تا کنون فرض بر این بود داده‌ها توسط یک خط جدا می‌شوند.



تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

ماشین بردار پشتیبان

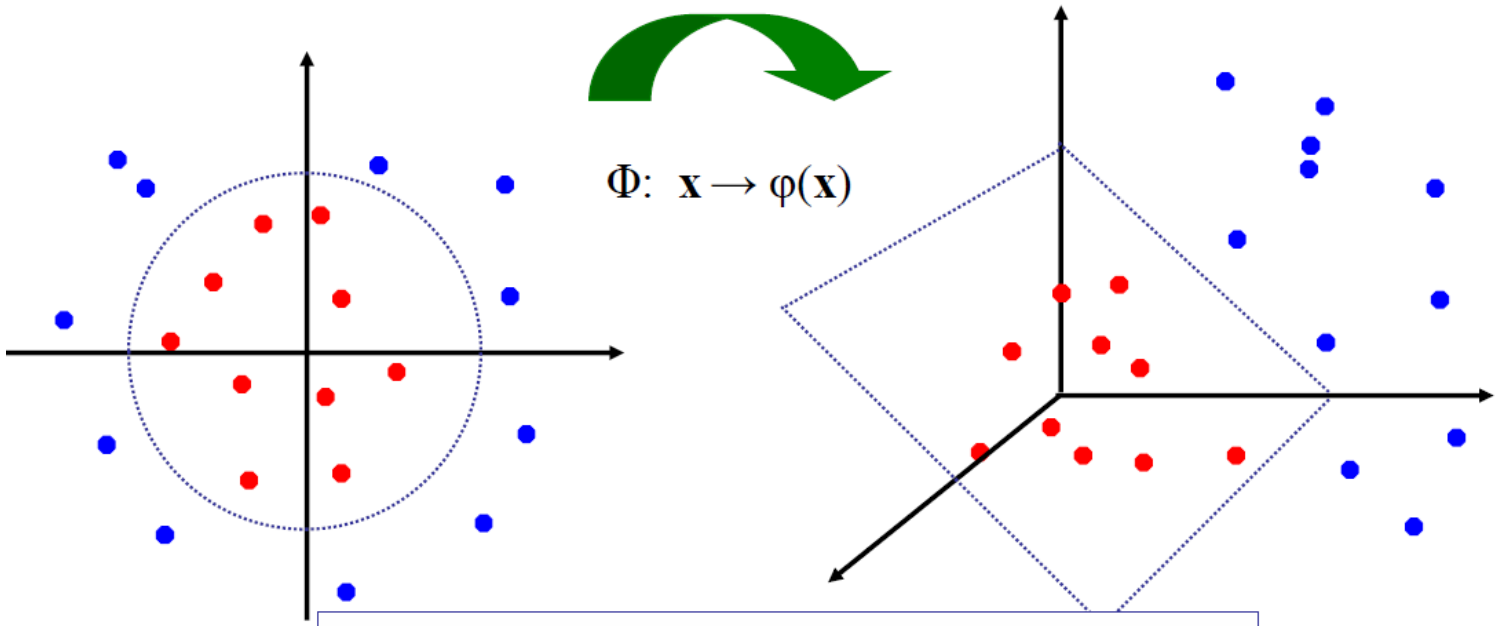
- ماشین بردار پشتیبان غیرخطی
- داده‌ها را به یک فضا با بعد بالا نگاشت می‌دهیم به امید آنکه تفکیک‌پذیری خطی حاصل شود.



تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

ماشین بردار پشتیبان غیرخطی

- ایده کلی – نگاشت داده‌ها به بعد بالاتر به امید حصول تفکیک‌پذیری خطی



$$\Phi : \mathbb{R}^2 \to \mathbb{R}^3$$
$$(x_1, x_2) \mapsto (z_1, z_2, z_3) := (x_1^2, \sqrt{2}\, x_1 x_2, x_2^2)$$

تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

ماشین بردار پشتیبان غیرخطی

$$W^* = \arg\min \frac{1}{2}|W|^2$$

$$y_n(w\phi(x_n) + b) \geq 1$$

$$W^* = \arg\min \left(\frac{1}{2}|W|^2 - \sum_{i=1}^{n} \alpha_i (y_i(w^T\phi(x_i) + b) - 1) \right)$$

$$J = \frac{1}{2}w^T w - \sum_{i=1}^{n} \alpha_i y_i w^T \phi(x_i) - b\sum_{i=1}^{n} \alpha_i y_i + \sum_{i=1}^{n} \alpha_i$$

$$w = \sum_{i=1}^{n} \alpha_i y_i \phi(x_i), \sum_{i=1}^{n} \alpha_i y_i = 0$$

$$J = \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j y_i y_j \underbrace{\phi(x_i)^T \phi(x_j)}_{K(x_i, x_j)}$$

$$b = \frac{1}{N_s}(\Sigma y_n - \sum \alpha_m y_m \phi(x_i)^T \phi(x_j))$$

$$\sum_{i=1}^{n} \alpha_i y_i = 0, \alpha_i > 0$$

تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

شرایط Mercer

- آیا تابع کرنل می‌تواند هر تابع دلخواهی باشد؟
  - کرنل یک تابع پیوسته و متقارن است.
  - در فضای اقلیدسی، یک کرنل به صورت زیر تعریف می‌شود:

$$K(x,x') = \sum_{i=1}^{\infty} \lambda_i \phi_i(x)^T \phi_i(x')$$

  - شرط Mercer

$$\int_a^b \int_a^b K(x,x')\psi(x)\psi(x')dxdx' \geq 0\,, \qquad \int_a^b \psi^2(x)dx \leq \infty$$

تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

ماشین بردار پشتیبان

- مسئله XOR

کلاس 1 $x_1 = [-1, -1], \quad x_4 = [1, 1]$

کلاس 2 $x_2 = [-1, 1], \quad x_3 = [1, -1]$

$$\Longrightarrow \quad K(x, x') = (x^T x' + 1)^2$$

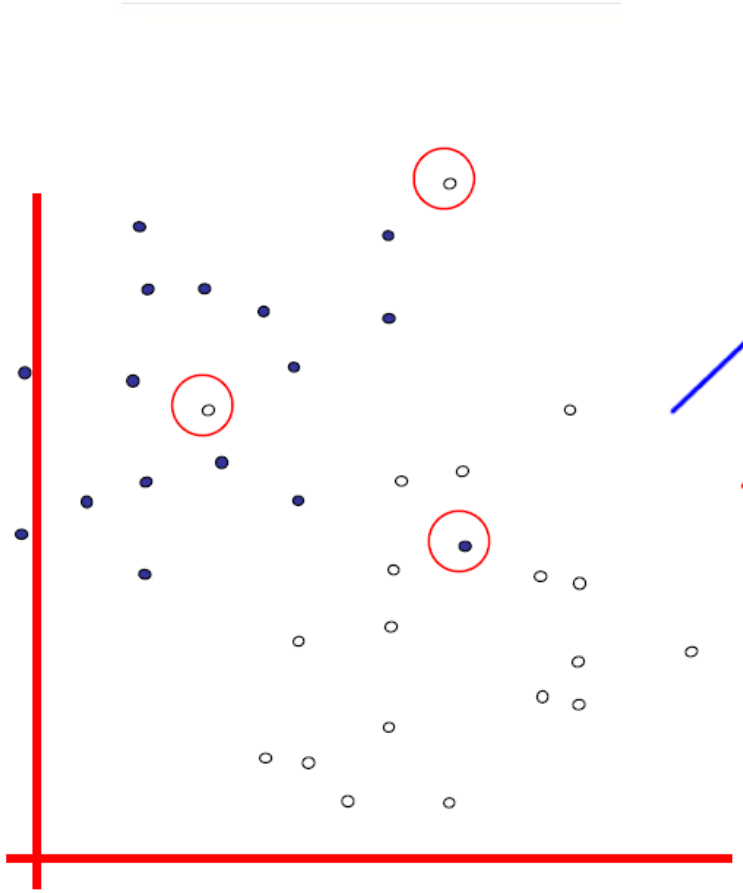
$$\phi(x) = [1 \quad \sqrt{2}x_1 \quad \sqrt{2}x_2 \quad \sqrt{2}x_1 x_2 \quad x_1^2 \quad x_2^2]$$

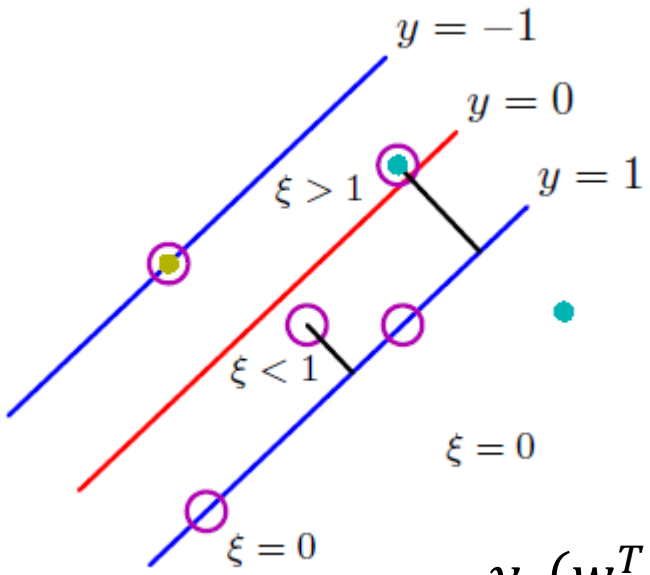
$$\alpha^* = \arg\min \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2}\sum_{i=1}^{4}\sum_{j=1}^{4} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \Longrightarrow \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \frac{1}{8}$$

$$g(x) = x_1 x_2$$

تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
dayche.com | گروه دایکه

ماشین بردار پشتیبان





- کلاس‌های جدایی‌ناپذیر خطی

کران بالای کلاس‌های گمشده

$$\arg\min C \sum_{i=1}^{N} \epsilon_i + \frac{1}{2} w^T w$$

$$y_n(w^T \phi(x_n) + b) \geq 1 - \epsilon_n \qquad \epsilon_n \geq 0$$

$$J = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j y_i y_j \phi(x_i)^T \phi(x_j)$$

$$\sum_{i=1}^{n} \alpha_i y_i = 0, 0 \leq \alpha_i \leq C$$

تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
گروه دایکه | dayche.com

ماشین بردار پشتیبان – مسئله رگرسیون

- مدل رگرسیون خطی

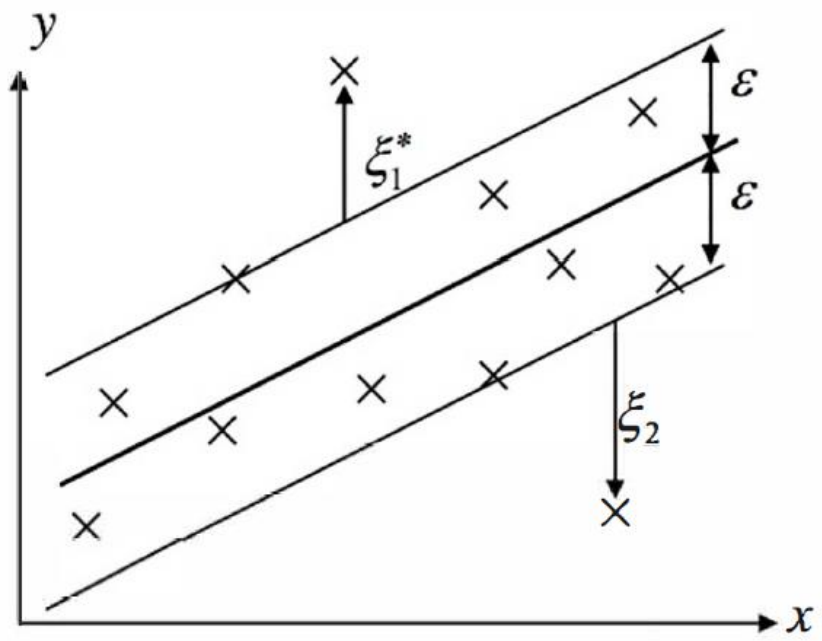
$$y = w^T x + w_0 \rightarrow W^* = \arg\min (y - \hat{y})^2 + \lambda W^T W$$

$$W^* = \arg\min \frac{1}{2} W^T W + C \sum_{i=1}^{n} (\epsilon_{i1} + \epsilon_{i2})$$

$$y - \hat{y} \leq \epsilon_1 + \epsilon$$

$$\hat{y} - y \leq \epsilon_2 + \epsilon$$

$$\epsilon_1, \epsilon_2 \geq 0$$



تولید محتوا: وحید محمدزاده ایوقی
daychegroup
daychegroup
گروه دایکه | dayche.com