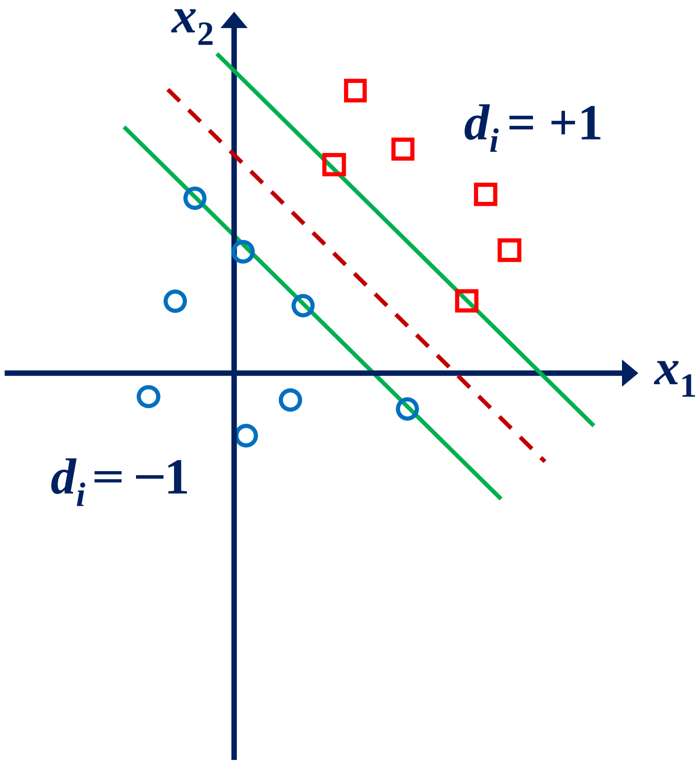
شبکه‌های عصبی مصنوعی

جلسه پانزدهم:
ماشین بردار پشتیبان (۳)
(Support Vector Machine = SVM)

محمد فرخی (farrokhi AT iust.ac.ir)

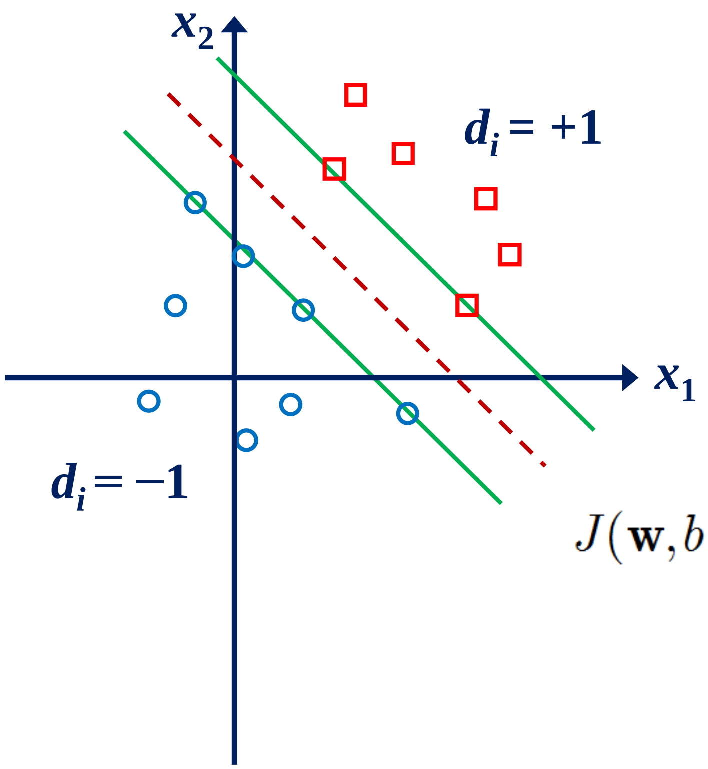
ماشین بردار پشتیبان (SVM)

۱- ابرصفحه بهینه برای الگوهای جداپذیر خطی



ماشین بردار پشتیبان (SVM)

۱- ابرصفحه بهینه برای الگوهای جداپذیر خطی



- حل مساله اولیه و یافتن وزن‌های بهینه:

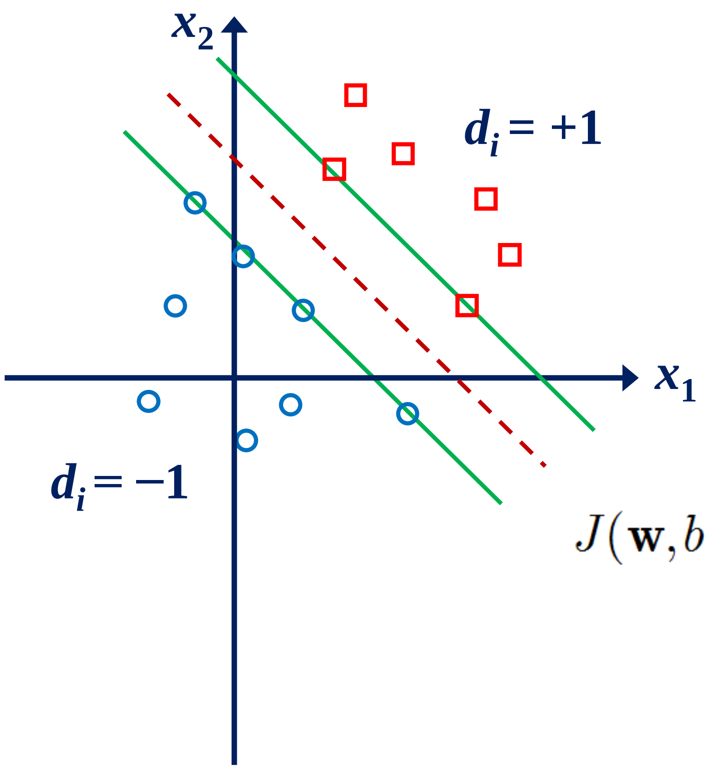
$$\phi(\mathbf{w}) = \min \frac{1}{2}\mathbf{w}^T\mathbf{w}$$

$$d_i(\mathbf{w}^T\mathbf{x}_i + b) - 1 = 0 \qquad \text{for} \qquad i = 1,\ldots,N$$

- تابع لاگرانژ

$$J(\mathbf{w}, b, \alpha) = \frac{1}{2}\mathbf{w}^T\mathbf{w} - \sum_{i=1}^{N} \alpha_i \left[d_i(\mathbf{w}^T\mathbf{x}_i + b) - 1 \right]$$

ماشین بردار پشتیبان (SVM)



۱- ابرصفحه بهینه برای الگوهای جداپذیر خطی

– حل مساله اولیه و یافتن وزن‌های بهینه:

$$\phi(\mathbf{w}) = \min \frac{1}{2}\mathbf{w}^T\mathbf{w}$$

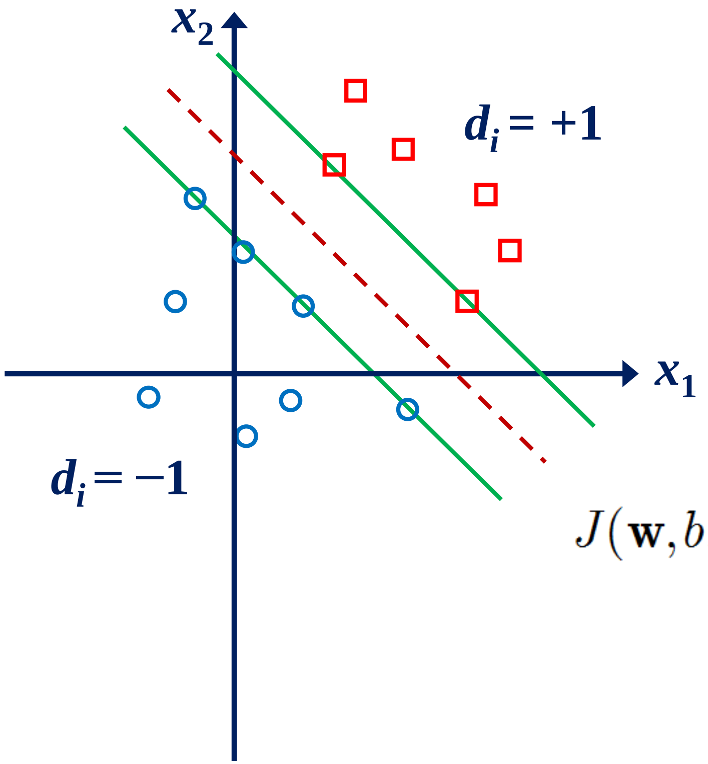
$$d_i(\mathbf{w}^T\mathbf{x}_i + b) - 1 = 0 \quad \text{for} \quad i = 1, \ldots, N$$

– تابع لاگرانژ

$$J(\mathbf{w}, b, \alpha) = \frac{1}{2}\mathbf{w}^T\mathbf{w} - \sum_{i=1}^{N} \alpha_i \left[d_i(\mathbf{w}^T\mathbf{x}_i + b) - 1\right]$$

ایجاد نقطه زینی (Saddle Point)

ماشین بردار پشتیبان (SVM)



۱– ابرصفحه بهینه برای الگوهای جداپذیر خطی

– حل مساله اولیه و یافتن وزن‌های بهینه:

$$\phi(\mathbf{w}) = \min \frac{1}{2}\mathbf{w}^T\mathbf{w}$$

$$d_i(\mathbf{w}^T\mathbf{x}_i + b) - 1 = 0 \qquad \text{for} \qquad i = 1,\dots,N$$

– تابع لاگرانژ

$$J(\mathbf{w}, b, \alpha) = \frac{1}{2}\mathbf{w}^T\mathbf{w} - \sum_{i=1}^{N}\alpha_i\left[d_i(\mathbf{w}^T\mathbf{x}_i + b) - 1\right]$$

ایجاد نقطه زینی (Saddle Point)

– حل مساله دوگان و یافتن ضرایب لاگرانژ:

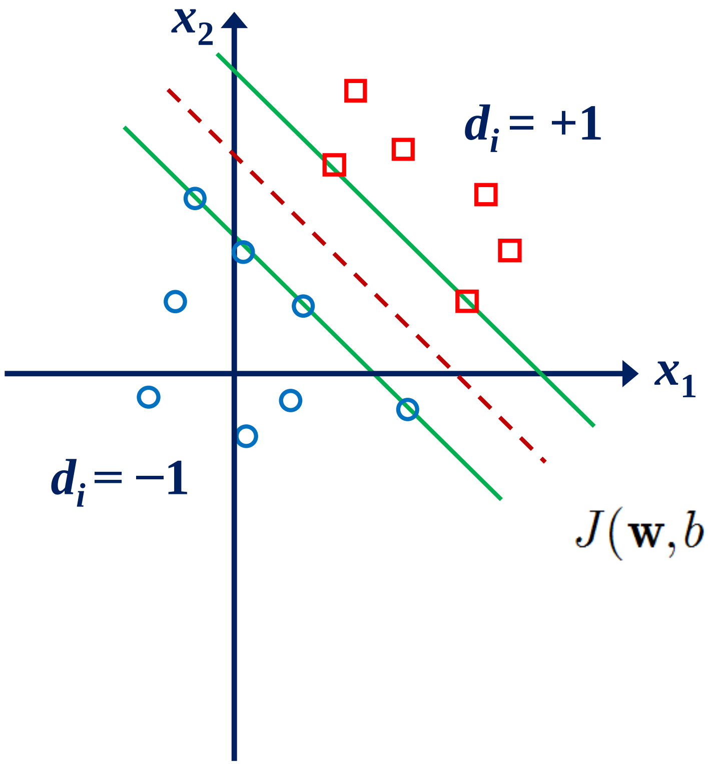
$$\phi(\alpha) = \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j d_i d_j \mathbf{x}_i^T\mathbf{x}_j$$

$$\sum_{i=1}^{N}\alpha_i d_i = 0$$

$$\alpha_i \geq 0 \qquad i = 1,\dots,N$$

ماشین بردار پشتیبان (SVM)

۱- ابرصفحه بهینه برای الگوهای جداپذیر خطی



– حل مساله اولیه و یافتن وزن‌های بهینه:

$$\phi(\mathbf{w}) = \min \frac{1}{2}\mathbf{w}^T\mathbf{w}$$

$$d_i(\mathbf{w}^T\mathbf{x}_i + b) - 1 = 0 \quad \text{for} \quad i = 1,\dots,N$$

– تابع لاگرانژ

$$J(\mathbf{w},b,\alpha) = \frac{1}{2}\mathbf{w}^T\mathbf{w} - \sum_{i=1}^{N}\alpha_i\left[d_i(\mathbf{w}^T\mathbf{x}_i + b) - 1\right]$$

ایجاد نقطه زینی (Saddle Point)

– حل مساله دوگان و یافتن ضرایب لاگرانژ:

$$\phi(\alpha) = \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j d_i d_j \mathbf{x}_i^T\mathbf{x}_j$$

$$\sum_{i=1}^{N}\alpha_i d_i = 0$$

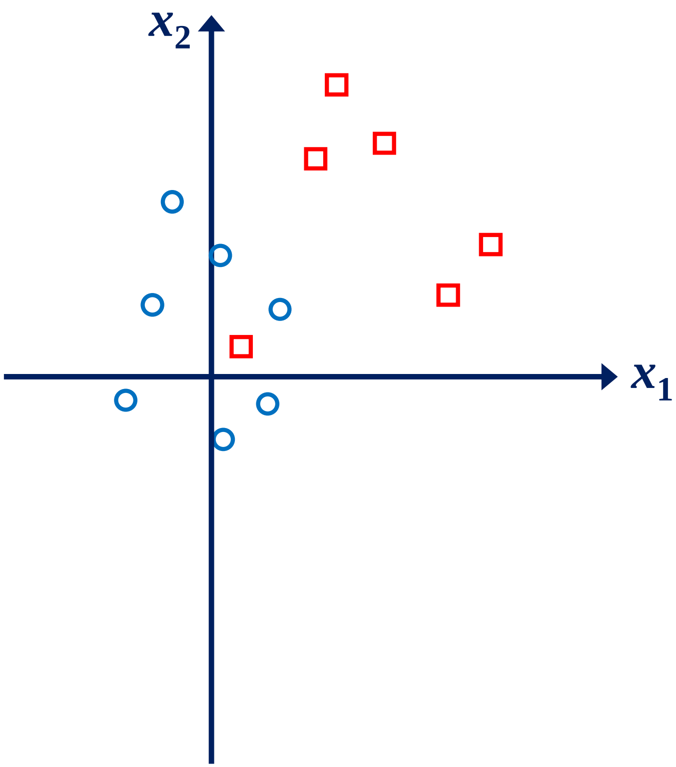
$$\alpha_i \geq 0 \qquad i = 1,\dots,N$$

تعیین بردارهای پشتیبان $\mathbf{x}^{(s)}$

ماشین بردار پشتیبان (SVM)

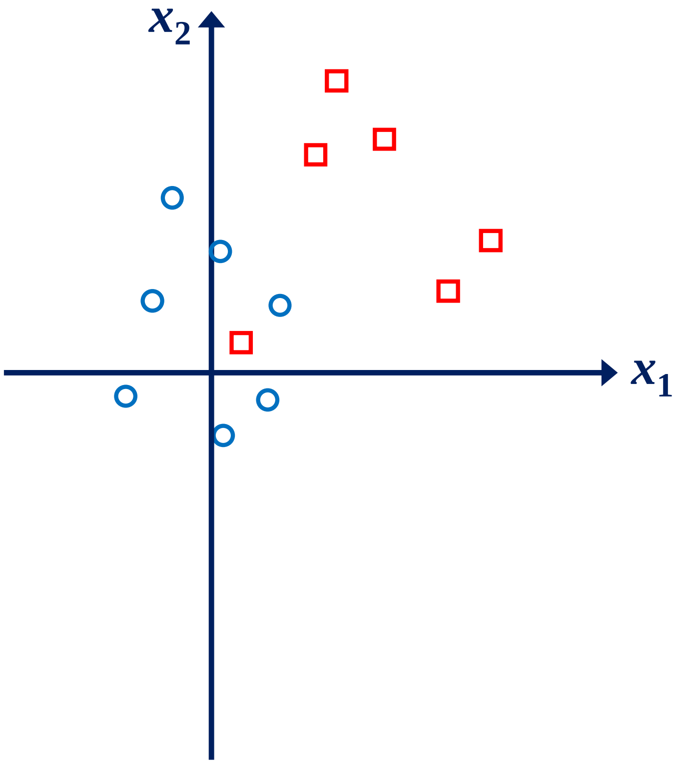
ماشین بردار پشتیبان (SVM)

۲- ابرصفحه بهینه برای الگوهای جداناپذیر



ماشین بردار پشتیبان (SVM)

۲- ابرصفحه بهینه برای الگوهای جداناپذیر



- حل مساله اولیه و یافتن وزن‌های بهینه:

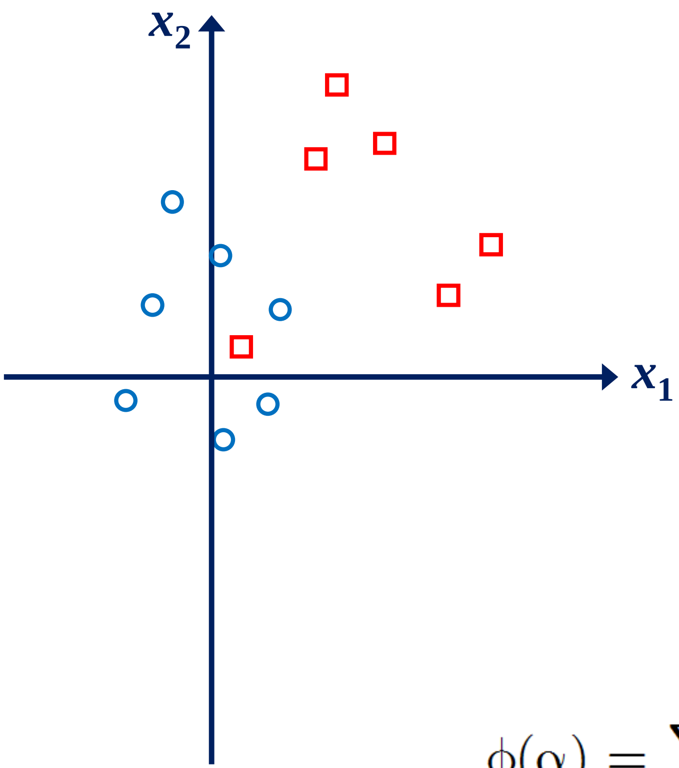
$$J(\mathbf{w}, b, \alpha, \xi) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{N}\xi_i - \sum_{i=1}^{N}\alpha_i\left[d_i(\mathbf{w}^T\mathbf{x}_i + b) - 1 + \xi_i\right]$$

$$d_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 - \xi_i \qquad i = 1, \ldots, N$$

$$\xi_i > 0 \quad \forall i$$

ماشین بردار پشتیبان (SVM)

۲- ابرصفحه بهینه برای الگوهای جداناپذیر



- حل مساله اولیه و یافتن وزن‌های بهینه:

$$J(\mathbf{w},b,\alpha,\xi)=\frac{1}{2}\mathbf{w}^T\mathbf{w}+C\sum_{i=1}^{N}\xi_i - \sum_{i=1}^{N}\alpha_i\left[d_i(\mathbf{w}^T\mathbf{x}_i+b)-1+\xi_i\right]$$

$$d_i(\mathbf{w}^T\mathbf{x}_i+b)\geq 1-\xi_i \qquad i=1,\dots,N$$

$$\xi_i>0 \quad \forall i$$

- حل مساله دوگان و یافتن ضرایب لاگرانژ:

$$\phi(\alpha)=\sum_{i=1}^{N}\alpha_i-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j d_i d_j \mathbf{x}_i^T\mathbf{x}_j$$

$$\sum_{i=1}^{N}\alpha_i d_i=0$$

$$0\leq\alpha_i\leq C$$

تعیین بردارهای پشتیبان $\mathbf{x}^{(s)}$

ماشین بردار پشتیبان (SVM)

– فلسفه SVM برای کلاسه‌بندی الگوها

ماشین بردار پشتیبان (SVM)

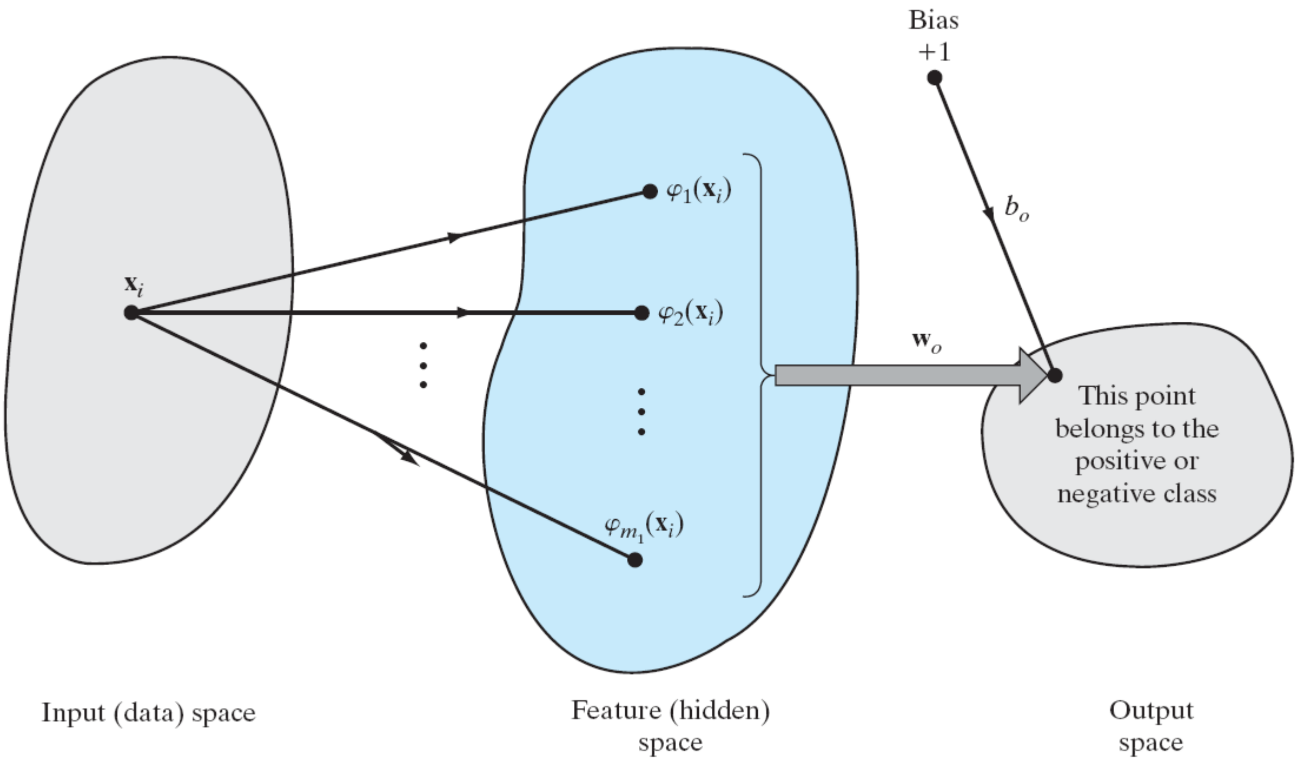
- فلسفه SVM برای کلاسه‌بندی الگوها
- اصول SVM

ماشین بردار پشتیبان (SVM)

- فلسفه SVM برای کلاسه‌بندی الگوها

- اصول SVM

۱- نگاشت غیرخطی بردار ورودی به فضای ویژگی‌ها

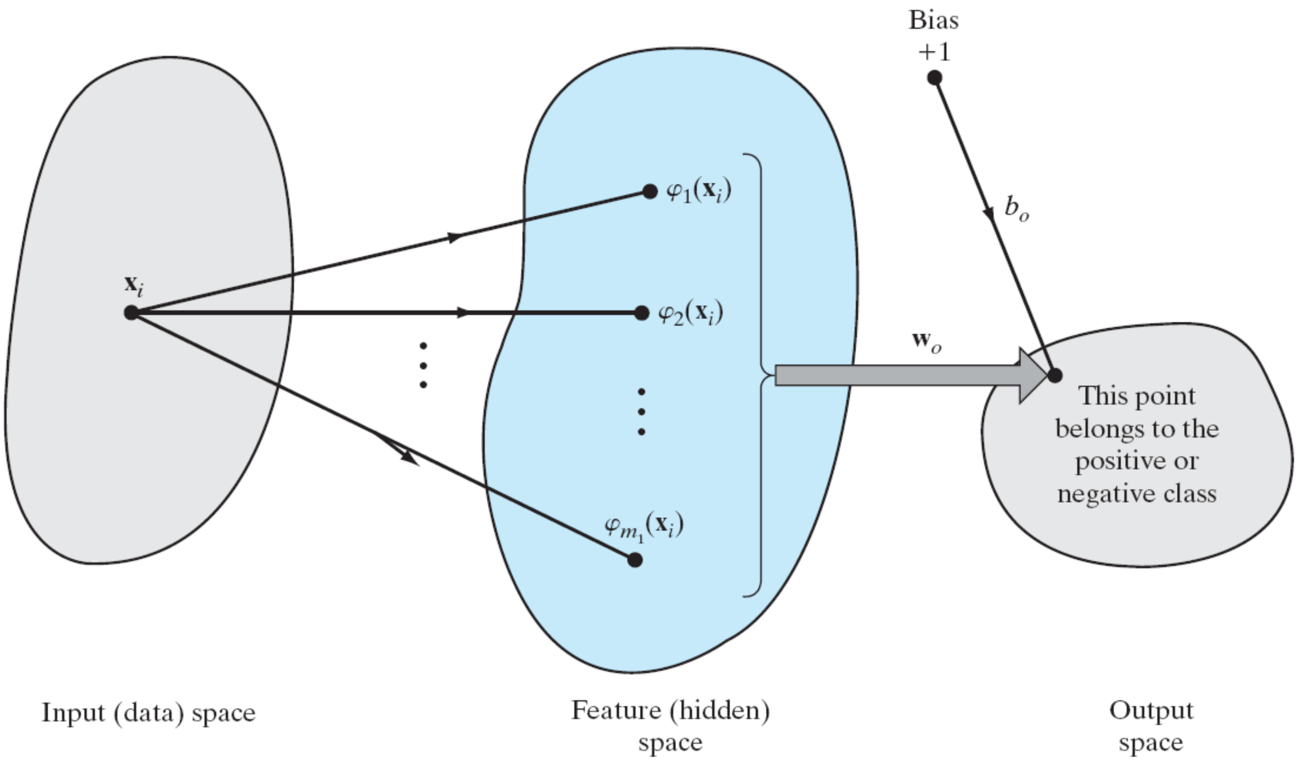


ماشین بردار پشتیبان (SVM)

- فلسفه SVM برای کلاسه‌بندی الگوها
- اصول SVM

۱- نگاشت غیرخطی بردار ورودی به فضای ویژگی‌ها

۲- تشکیل ابرصفحه بهینه برای جداسازی ویژگی‌ها که در قسمت ۱ کشف شدند.



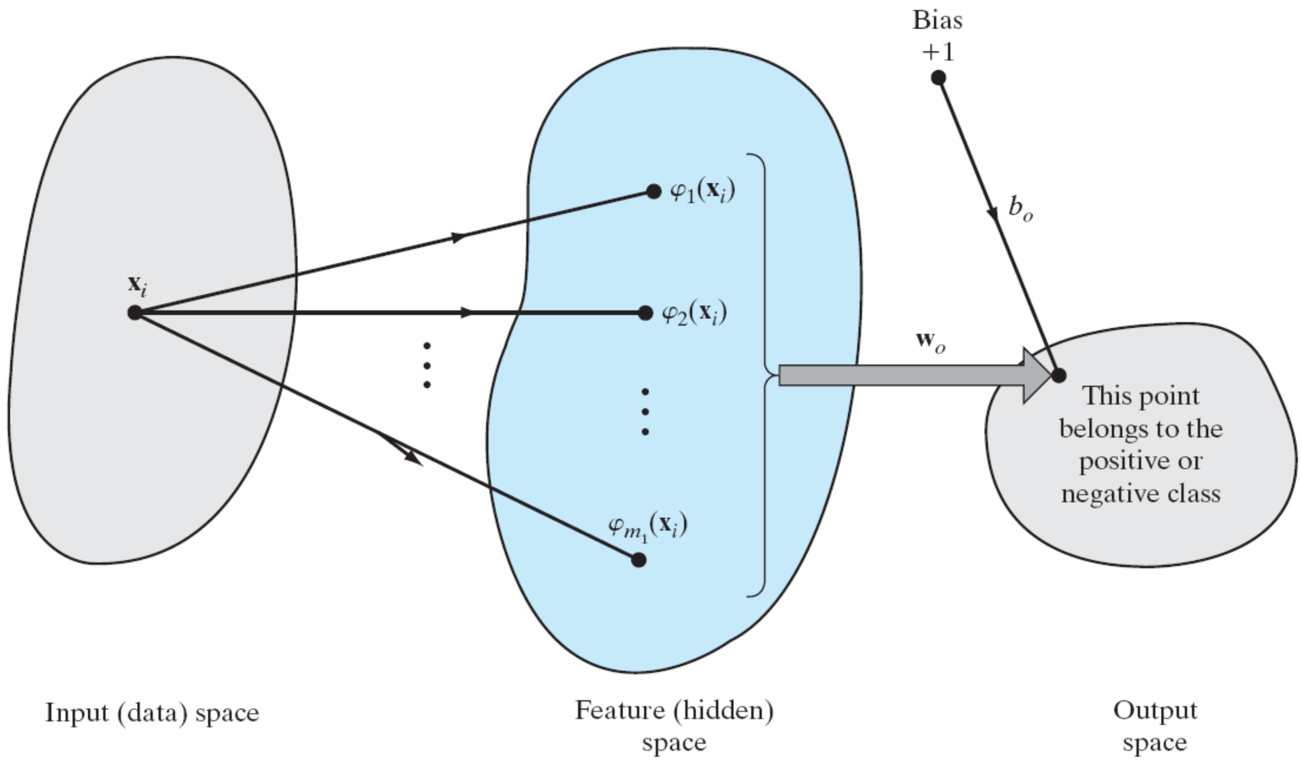
ماشین بردار پشتیبان (SVM)

- فلسفه SVM برای کلاسه‌بندی الگوها

- اصول SVM

۱- نگاشت غیرخطی بردار ورودی به فضای ویژگی‌ها

۲- تشکیل ابرصفحه بهینه برای جداسازی ویژگی‌ها که در قسمت ۱ کشف شدند.



نکته مهم: نظریه SVM، به‌طور تحلیلی ابعاد بهینه برای فضای پنهان پیدامی‌کند.

ماشین بردار پشتیبان (SVM)

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

- فرض کنید:

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

- فرض‌کنید:
  - بردار x از فضای ورودی با بُعد m_0 انتخاب شده است

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

– فرض‌کنید:

- بردار $\mathbf{x}$ از فضای ورودی با بُعد m_0 انتخاب شده است
- مجموعه‌ای از توابع غیرخطی $\{\varphi_j(\mathbf{x})\}_{j=1}^{\infty}$ که فضای ورودی را به فضای ویژگی‌ها انتقال می‌دهد.

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

- فرض‌کنید:
  - بردار $\mathbf{x}$ از فضای ورودی با بُعد m_0 انتخاب شده است
  - $\{\varphi_j(\mathbf{x})\}_{j=1}^{\infty}$ مجموعه‌ای از توابع غیرخطی که فضای ورودی را به فضای ویژگی‌ها انتقال می‌دهد.

- ابرصفحه به‌عنوان سطح تصمیم‌گیری

$$\sum_{j=1}^{\infty} w_j \varphi_j(\mathbf{x}) = 0$$

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

– فرض‌کنید:

- بردار $\mathbf{x}$ از فضای ورودی با بُعد m_0 انتخاب شده است
- $\{\varphi_j(\mathbf{x})\}_{j=1}^{\infty}$ مجموعه‌ای از توابع غیرخطی که فضای ورودی را به فضای ویژگی‌ها انتقال می‌دهد.

– ابرصفحه به‌عنوان سطح تصمیم‌گیری

$$\sum_{j=1}^{\infty} w_j \varphi_j(\mathbf{x}) = 0$$

$\{\mathbf{w}_j\}_{j=1}^{\infty}$ مجموعه‌ای بی‌نهایت بزرگ از وزن‌ها که فضای ویژگی‌ها را به فضای خروجی انتقال می‌دهد.

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

– فرض‌کنید:

- بردار $\mathbf{x}$ از فضای ورودی با بُعد m_0 انتخاب شده است
- $\{\varphi_j(\mathbf{x})\}_{j=1}^{\infty}$ مجموعه‌ای از توابع غیرخطی که فضای ورودی را به فضای ویژگی‌ها انتقال می‌دهد.

– ابرصفحه به‌عنوان سطح تصمیم‌گیری

$$\sum_{j=1}^{\infty} w_j \varphi_j(\mathbf{x}) = 0$$

$\{\mathbf{w}_j\}_{j=1}^{\infty}$ مجموعه‌ای بی‌نهایت بزرگ از وزن‌ها که فضای ویژگی‌ها را به فضای خروجی انتقال می‌دهد.

– به فرم برداری

$$\mathbf{w}^T \Phi(\mathbf{x}) = 0$$

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

– فرض‌کنید:

- بردار $\mathbf{x}$ از فضای ورودی با بُعد m_0 انتخاب شده است
- $\{\varphi_j(\mathbf{x})\}_{j=1}^{\infty}$ مجموعه‌ای از توابع غیرخطی که فضای ورودی را به فضای ویژگی‌ها انتقال می‌دهد.

– ابرصفحه به‌عنوان سطح تصمیم‌گیری

$$\sum_{j=1}^{\infty} w_j \varphi_j(\mathbf{x}) = 0$$

$\{\mathbf{w}_j\}_{j=1}^{\infty}$ مجموعه‌ای بی‌نهایت بزرگ از وزن‌ها که فضای ویژگی‌ها را به فضای خروجی انتقال می‌دهد.

– به فرم برداری

$$\mathbf{w}^T \Phi(\mathbf{x}) = 0$$

$\Phi(\mathbf{x})$ بردار ویژگی‌ها

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

- همانند قبل، به‌دنبال «جداپذیری خطی» الگوهای انتقال‌یافته به فضای ویژگی‌ها هستیم.

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

- همانند قبل، به‌دنبال «جداپذیری خطی» الگوهای انتقال‌یافته به فضای ویژگی‌ها هستیم.
- با توجه به این هدف، معادله وزن‌ها را که قبلا به‌صورت $\mathbf{w} = \sum_{i=1}^{Ns} \alpha_i d_i \mathbf{x}_i$ بود، را به‌فرم زیر ارایه می‌کنیم:

$$\mathbf{w} = \sum_{i=1}^{Ns} \alpha_i d_i \Phi(\mathbf{x}_i)$$

$$\Phi(\mathbf{x}_i) = [\varphi_1(\mathbf{x}_i), \varphi_2(\mathbf{x}_i), \ldots]^T$$

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

- همانند قبل، به‌دنبال «جداپذیری خطی» الگوهای انتقال‌یافته به فضای ویژگی‌ها هستیم.
- با توجه به این هدف، معادله وزن‌ها را که قبلا به‌صورت $\mathbf{w} = \sum_{i=1}^{Ns} \alpha_i d_i \mathbf{x}_i$ بود، را به‌فرم زیر ارایه می‌کنیم:

$$\mathbf{w} = \sum_{i=1}^{Ns} \alpha_i d_i \Phi(\mathbf{x}_i)$$

$$\Phi(\mathbf{x}_i) = [\varphi_1(\mathbf{x}_i), \varphi_2(\mathbf{x}_i), \ldots]^T$$

- با قراردادن این رابطه در $\mathbf{w}^T \Phi(\mathbf{x}) = 0$ خواهیم داشت:

$$\sum_{i=1}^{Ns} \alpha_i d_i \Phi^T(\mathbf{x}_i)\, \Phi(\mathbf{x}) = 0$$

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

- همانند قبل، به‌دنبال «جداپذیری خطی» الگوهای انتقال‌یافته به فضای ویژگی‌ها هستیم.

- با توجه به این هدف، معادله وزن‌ها را که قبلا به‌صورت $\mathbf{w} = \sum_{i=1}^{Ns} \alpha_i d_i \mathbf{x}_i$ بود، را به‌فرم زیر ارایه می‌کنیم:

$$\mathbf{w} = \sum_{i=1}^{Ns} \alpha_i d_i \Phi(\mathbf{x}_i)$$

$$\Phi(\mathbf{x}_i) = [\varphi_1(\mathbf{x}_i), \varphi_2(\mathbf{x}_i), \ldots]^T$$

- با قراردادن این رابطه در $\mathbf{w}^T \Phi(\mathbf{x}) = 0$ خواهیم داشت:

$$\sum_{i=1}^{Ns} \alpha_i d_i \Phi^T(\mathbf{x}_i) \Phi(\mathbf{x}) = 0$$

- ضرب داخلی را با اسکالر زیر نشان می‌دهیم:

$$k(\mathbf{x}, \mathbf{x}_i) = \Phi^T(\mathbf{x}_i) \Phi(\mathbf{x})$$

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

– همانند قبل، به‌دنبال «جداپذیری خطی» الگوهای انتقال‌یافته به فضای ویژگی‌ها هستیم.

– با توجه به این هدف، معادله وزن‌ها را که قبلا به‌صورت $\mathbf{w} = \sum_{i=1}^{Ns} \alpha_i d_i \mathbf{x}_i$ بود، را به‌فرم زیر ارایه می‌کنیم:

$$\mathbf{w} = \sum_{i=1}^{Ns} \alpha_i d_i \Phi(\mathbf{x}_i)$$

$$\Phi(\mathbf{x}_i) = [\varphi_1(\mathbf{x}_i), \varphi_2(\mathbf{x}_i), \dots]^T$$

– با قراردادن این رابطه در $\mathbf{w}^T \Phi(\mathbf{x}) = 0$ خواهیم داشت:

$$\sum_{i=1}^{Ns} \alpha_i d_i \Phi^T(\mathbf{x}_i)\, \Phi(\mathbf{x}) = 0$$

– ضرب داخلی را با اسکالر زیر نشان می‌دهیم:

$$\begin{aligned} k(\mathbf{x}, \mathbf{x}_i) &= \Phi^T(\mathbf{x}_i)\, \Phi(\mathbf{x}) \\ &= \sum_{j=1}^{\infty} \varphi_j(\mathbf{x}_i)\, \varphi_j(\mathbf{x}) \qquad i = 1, \dots, Ns \end{aligned}$$

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

- بنابراین، سطح تصمیم‌گیری بهینه (ابرصفحه) در فضای خروجی برابر است با

$$\sum_{i=1}^{Ns} \alpha_i d_i k(\mathbf{x}, \mathbf{x}_i) = 0$$

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

- بنابراین، سطح تصمیم‌گیری بهینه (ابرصفحه) در فضای خروجی برابر است با

$$\sum_{i=1}^{Ns} \alpha_i d_i k(\mathbf{x}, \mathbf{x}_i) = 0$$

$k(\mathbf{x}, \mathbf{x}_i)$ ضرب داخلی کرنل (یا به‌طور ساده، کرنل) که به‌صورت زیر تعریف می‌شود:

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

- بنابراین، سطح تصمیم‌گیری بهینه (ابرصفحه) در فضای خروجی برابر است با

$$\sum_{i=1}^{Ns} \alpha_i d_i k(\mathbf{x},\mathbf{x}_i) = 0$$

$k(\mathbf{x},\mathbf{x}_i)$ ضرب داخلی کرنل (یا به‌طور ساده، کرنل) که به‌صورت زیر تعریف می‌شود:

کرنل $k(\mathbf{x},\mathbf{x}_i)$ تابعی است که ضرب داخلی تصاویر تشکیل‌شده در فضای ویژگی‌ها را برای دو نقطه در فضای ورودی با استفاده از $\Phi($ محاسبه می‌کند.

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

- بنابراین، سطح تصمیم‌گیری بهینه (ابرصفحه) در فضای خروجی برابر است با

$$\sum_{i=1}^{Ns} \alpha_i d_i k(\mathbf{x}, \mathbf{x}_i) = 0$$

$k(\mathbf{x}, \mathbf{x}_i)$ ضرب داخلی کرنل (یا به‌طور ساده، کرنل) که به‌صورت زیر تعریف می‌شود:

کرنل $k(\mathbf{x}, \mathbf{x}_i)$ تابعی است که ضرب داخلی تصاویر تشکیل‌شده در فضای ویژگی‌ها را برای دو نقطه در فضای ورودی با استفاده از $\Phi($ محاسبه می‌کند.

- دوخاصیت کرنل:

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

- بنابراین، سطح تصمیم‌گیری بهینه (ابرصفحه) در فضای خروجی برابر است با

$$\sum_{i=1}^{Ns} \alpha_i d_i k(\mathbf{x}, \mathbf{x}_i) = 0$$

$k(\mathbf{x}, \mathbf{x}_i)$ ضرب داخلی کرنل (یا به‌طور ساده، کرنل) که به‌صورت زیر تعریف می‌شود:

کرنل $k(\mathbf{x}, \mathbf{x}_i)$ تابعی است که ضرب داخلی تصاویر تشکیل‌شده در فضای ویژگی‌ها را برای دو نقطه در فضای ورودی با استفاده از Φ محاسبه می‌کند.

- دوخاصیت کرنل:

۱- تابع $k(\mathbf{x}, \mathbf{x}_i)$ حول نقطه مرکز ($\mathbf{x}_i$) متقارن است

$$k(\mathbf{x}, \mathbf{x}_i) = k(\mathbf{x}_i, \mathbf{x}) \quad \forall\, \mathbf{x}_i$$

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

- بنابراین، سطح تصمیم‌گیری بهینه (ابرصفحه) در فضای خروجی برابر است با

$$\sum_{i=1}^{Ns} \alpha_i d_i k(\mathbf{x}, \mathbf{x}_i) = 0$$

$k(\mathbf{x}, \mathbf{x}_i)$ ضرب داخلی کرنل (یا به‌طور ساده، کرنل) که به‌صورت زیر تعریف می‌شود:

کرنل $k(\mathbf{x}, \mathbf{x}_i)$ تابعی است که ضرب داخلی تصاویر تشکیل‌شده در فضای ویژگی‌ها را برای دو نقطه در فضای ورودی با استفاده از Φ محاسبه می‌کند.

- دوخاصیت کرنل:

۱- تابع $k(\mathbf{x}, \mathbf{x}_i)$ حول نقطه مرکز ($\mathbf{x}_i$) متقارن است

$$k(\mathbf{x}, \mathbf{x}_i) = k(\mathbf{x}_i, \mathbf{x}) \qquad \forall \mathbf{x}_i$$

۲- سطح (حجم) زیر منحنی تابع $k(\mathbf{x}, \mathbf{x}_i)$ مقداری ثابت است.

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

شعبده بازی کرنل:

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

شعبده بازی کرنل:

– معادله سطح تصمیم‌گیری را مجددا درنظربگیرید:

$$\sum_{i=1}^{Ns} \alpha_i d_i k(\mathbf{x}, \mathbf{x}_i) = 0$$

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

شعبده بازی کرنل:

- معادله سطح تصمیم‌گیری را مجددا درنظربگیرید:

$$\sum_{i=1}^{Ns} \alpha_i d_i k(\mathbf{x}, \mathbf{x}_i) = 0$$

- در مورد این معادله، دو نکته قابل تامل است:

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

شعبده بازی کرنل:

– معادله سطح تصمیم‌گیری را مجددا درنظربگیرید:

$$\sum_{i=1}^{Ns} \alpha_i d_i k(\mathbf{x}, \mathbf{x}_i) = 0$$

– در مورد این معادله، دو نکته قابل تامل است:

۱- برای کلاسه‌بندی الگوها در فضای خروجی، فقط نیاز به کرنل $k(\mathbf{x}, \mathbf{x}_i)$ داریم.

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

شعبده بازی کرنل:

– معادله سطح تصمیم‌گیری را مجددا درنظربگیرید:

$$\sum_{i=1}^{Ns} \alpha_i d_i k(\mathbf{x}, \mathbf{x}_i) = 0$$

– در مورد این معادله، دو نکته قابل تامل است:

۱- برای کلاسه‌بندی الگوها در فضای خروجی، فقط نیاز به کرنل $k(\mathbf{x}, \mathbf{x}_i)$ داریم.

به‌عبارت دیگر، هیچگاه نیاز به محاسبه صریح بردار وزن $\mathbf{w}_o$ نداریم.

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

شعبده بازی کرنل:

- معادله سطح تصمیم‌گیری را مجددا درنظربگیرید:

$$\sum_{i=1}^{Ns} \alpha_i d_i k(\mathbf{x}, \mathbf{x}_i) = 0$$

- در مورد این معادله، دو نکته قابل تامل است:

۱- برای کلاسه‌بندی الگوها در فضای خروجی، فقط نیاز به کرنل $k(\mathbf{x}, \mathbf{x}_i)$ داریم.

به‌عبارت دیگر، هیچگاه نیاز به محاسبه صریح بردار وزن $\mathbf{w}_o$ نداریم.

به‌همین جهت، استفاده از $k(\mathbf{x}, \mathbf{x}_i) = \Phi^T(\mathbf{x}_i)\Phi(\mathbf{x})$ را شعبده‌بازی کرنل می‌نامند.

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

شعبده بازی کرنل:

– معادله سطح تصمیم‌گیری را مجددا درنظربگیرید:

$$\sum_{i=1}^{Ns} \alpha_i d_i k(\mathbf{x}, \mathbf{x}_i) = 0$$

– در مورد این معادله، دو نکته قابل تامل است:

۱- برای کلاسه‌بندی الگوها در فضای خروجی، فقط نیاز به کرنل $k(\mathbf{x}, \mathbf{x}_i)$ داریم.
به‌عبارت دیگر، هیچگاه نیاز به محاسبه صریح بردار وزن $\mathbf{w}_o$ نداریم.
به‌همین جهت، استفاده از $k(\mathbf{x}, \mathbf{x}_i) = \Phi^T(\mathbf{x}_i)\Phi(\mathbf{x})$ را شعبده‌بازی کرنل می‌نامند.

۲- اگرچه فرض کردیم که فضای ویژگی‌ها می‌تواند دارای ابعاد بی‌نهایت باشد،
معادله خطی ابر صفحه بالا از تعداد محدودی جمله تشکیل می‌شود.

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

شعبده بازی کرنل:

- معادله سطح تصمیم‌گیری را مجددا درنظربگیرید:

$$\sum_{i=1}^{Ns} \alpha_i d_i k(\mathbf{x}, \mathbf{x}_i) = 0$$

- در مورد این معادله، دو نکته قابل تامل است:

۱- برای کلاسه‌بندی الگوها در فضای خروجی، فقط نیاز به کرنل $k(\mathbf{x}, \mathbf{x}_i)$ داریم.
به‌عبارت دیگر، هیچگاه نیاز به محاسبه صریح بردار وزن $\mathbf{w}_o$ نداریم.
به‌همین جهت، استفاده از $k(\mathbf{x}, \mathbf{x}_i) = \Phi^T(\mathbf{x}_i)\,\Phi(\mathbf{x})$ را شعبده‌بازی کرنل می‌نامند.

۲- اگرچه فرض کردیم که فضای ویژگی‌ها می‌تواند دارای ابعاد بی‌نهایت باشد،
معادله خطی ابر صفحه بالا از تعداد محدودی جمله تشکیل می‌شود.

- به خاطر نکته ۱، به ماشین بردار پشتیبان، ماشین کرنل نیز می‌گویند.

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

شعبده بازی کرنل:

- معادله سطح تصمیم‌گیری را مجددا درنظربگیرید:

$$\sum_{i=1}^{Ns} \alpha_i d_i k(\mathbf{x}, \mathbf{x}_i) = 0$$

- در مورد این معادله، دو نکته قابل تامل است:

۱- برای کلاسه‌بندی الگوها در فضای خروجی، فقط نیاز به کرنل $k(\mathbf{x}, \mathbf{x}_i)$ داریم.
به‌عبارت دیگر، هیچگاه نیاز به محاسبه صریح بردار وزن $\mathbf{w}_o$ نداریم.
به‌همین جهت، استفاده از $k(\mathbf{x}, \mathbf{x}_i) = \Phi^T(\mathbf{x}_i)\Phi(\mathbf{x})$ را شعبده‌بازی کرنل می‌نامند.

۲- اگرچه فرض کردیم که فضای ویژگی‌ها می‌تواند دارای ابعاد بی‌نهایت باشد،
معادله خطی ابر صفحه بالا از تعداد محدودی جمله تشکیل می‌شود.

- به خاطر نکته ۱، به ماشین بردار پشتیبان، ماشین کرنل نیز می‌گویند.
- برای مقاصد کلاسه‌بندی الگو، پارامترهای ماشین توسط یک بردار N-بُعدی تعیین می‌شود که برای جمله iام آن برابر است با $\alpha_i d_i \quad i = 1, \ldots, N$

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

– می‌توان $\{k(\mathbf{x}_i, \mathbf{x}_j)\}_{i,j=1}^{N}$ را به‌عنوان درایه ijام ماتریس متقارن زیر درنظرگرفت:

$$\mathbf{K} = \{k(\mathbf{x}_i, \mathbf{x}_j)\}_{i,j=1}^{N}$$

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

– می‌توان $\{k(\mathbf{x}_i, \mathbf{x}_j)\}_{i,j=1}^{N}$ را به‌عنوان درایه ijام ماتریس متقارن زیر درنظرگرفت:

$$\mathbf{K} = \{k(\mathbf{x}_i, \mathbf{x}_j)\}_{i,j=1}^{N}$$

ماتریس $\mathbf{K}$ ماتریسی غیرمنفی است به نام ماتریس کرنل که گرام (Gram) نیز می‌نامند.

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

- می‌توان $\{k(\mathbf{x}_i,\mathbf{x}_j)\}_{i,j=1}^{N}$ را به‌عنوان درایه ijام ماتریس متقارن زیر درنظرگرفت:

$$\mathbf{K} = \{k(\mathbf{x}_i,\mathbf{x}_j)\}_{i,j=1}^{N}$$

ماتریس $\mathbf{K}$ ماتریسی غیرمنفی است به نام ماتریس کرنل که گرام (Gram) نیز می‌نامند.

قضیه مرسر (Mercer's Theorem):

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

– می‌توان $\{k(\mathbf{x}_i,\mathbf{x}_j)\}_{i,j=1}^{N}$ را به‌عنوان درایه ijام ماتریس متقارن زیر درنظرگرفت:

$$\mathbf{K}=\{k(\mathbf{x}_i,\mathbf{x}_j)\}_{i,j=1}^{N}$$

ماتریس **K** ماتریسی غیرمنفی است به نام ماتریس کرنل که گرام (Gram) نیز می‌نامند.

<u>قضیه مرسر (Mercer’s Theorem):</u>

– فرض کنید $k(\mathbf{x},\mathbf{x}')$ کرنل متقارن و پیوسته‌ای باشد که در بازه $a\leq \mathbf{x},\mathbf{x}'\leq b$ تعریف شده باشد. بسط کرنل $k(\mathbf{x},\mathbf{x}')$ را می‌توان با استفاده از ثابت $\lambda_i>0$ به‌صورت زیر ارایه کرد:

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

– می‌توان $\{k(\mathbf{x}_i, \mathbf{x}_j)\}_{i,j=1}^{N}$ را به‌عنوان درایه ijام ماتریس متقارن زیر درنظرگرفت:

$$\mathbf{K} = \{k(\mathbf{x}_i, \mathbf{x}_j)\}_{i,j=1}^{N}$$

ماتریس **K** ماتریسی غیرمنفی است به نام ماتریس کرنل که گرام (Gram) نیز می‌نامند.

<u>قضیه مرسر (Mercer's Theorem):</u>

– فرض کنید $k(\mathbf{x}, \mathbf{x}')$ کرنل متقارن و پیوسته‌ای باشد که در بازه $a \le \mathbf{x}, \mathbf{x}' \le b$ تعریف شده باشد. بسط کرنل $k(\mathbf{x}, \mathbf{x}')$ را می‌توان با استفاده از ثابت $\lambda_i > 0$ به‌صورت زیر ارایه کرد:

$$k(\mathbf{x}, \mathbf{x}') = \sum_{i=1}^{\infty} \lambda_i \varphi_i(\mathbf{x}) \varphi_i(\mathbf{x}')$$

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

- می‌توان $\{k(\mathbf{x}_i,\mathbf{x}_j)\}_{i,j=1}^{N}$ را به‌عنوان درایه ijام ماتریس متقارن زیر درنظرگرفت:

$$\mathbf{K}=\{k(\mathbf{x}_i,\mathbf{x}_j)\}_{i,j=1}^{N}$$

ماتریس $\mathbf{K}$ ماتریسی غیرمنفی است به نام ماتریس کرنل که گرام (Gram) نیز می‌نامند.

<u>قضیه مرسر (Mercer's Theorem):</u>

- فرض کنید $k(\mathbf{x},\mathbf{x}')$ کرنل متقارن و پیوسته‌ای باشد که در بازه $a\leq\mathbf{x},\mathbf{x}'\leq b$ تعریف شده باشد. بسط کرنل $k(\mathbf{x},\mathbf{x}')$ را می‌توان با استفاده از ثابت $\lambda_i>0$ به‌صورت زیر ارایه کرد:

$$k(\mathbf{x},\mathbf{x}')=\sum_{i=1}^{\infty}\lambda_i\varphi_i(\mathbf{x})\varphi_i(\mathbf{x}')$$

برای این که این بسط معتبر باشد و همچنین همگرای مطلق و پیوسته باشد، شرط لازم و کافی زیر باید برقرار باشد:

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

- می‌توان $\{k(\mathbf{x}_i,\mathbf{x}_j)\}_{i,j=1}^{N}$ را به‌عنوان درایه ijام ماتریس متقارن زیر درنظرگرفت:

$$\mathbf{K} = \{k(\mathbf{x}_i,\mathbf{x}_j)\}_{i,j=1}^{N}$$

ماتریس **K** ماتریسی غیرمنفی است به نام ماتریس کرنل که گرام (Gram) نیز می‌نامند.

<u>قضیه مرسر (Mercer's Theorem):</u>

- فرض کنید $k(\mathbf{x},\mathbf{x}')$ کرنل متقارن و پیوسته‌ای باشد که در بازه $a \le \mathbf{x},\mathbf{x}' \le b$ تعریف شده باشد. بسط کرنل $k(\mathbf{x},\mathbf{x}')$ را می‌توان با استفاده از ثابت $\lambda_i > 0$ به‌صورت زیر ارایه کرد:

$$k(\mathbf{x},\mathbf{x}') = \sum_{i=1}^{\infty} \lambda_i \varphi_i(\mathbf{x}) \varphi_i(\mathbf{x}')$$

برای این که این بسط معتبر باشد و همچنین همگرای مطلق و پیوسته باشد، شرط لازم و کافی زیر باید برقرار باشد:

$$\int_a^b \int_a^b k(\mathbf{x},\mathbf{x}') \psi(\mathbf{x}) \psi(\mathbf{x}')\, d\mathbf{x}\, d\mathbf{x}' = 0$$

به‌طوری‌که

$$\int_a^b \psi^2(\mathbf{x})\, d\mathbf{x} < \infty$$

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

$$k(\mathbf{x},\mathbf{x}') = \sum_{i=1}^{\infty} \lambda_i \varphi_i(\mathbf{x}) \varphi_i(\mathbf{x}')$$

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

$$k(\mathbf{x},\mathbf{x}') = \sum_{i=1}^{\infty} \lambda_i \varphi_i(\mathbf{x}) \varphi_i(\mathbf{x}')$$

$\varphi_i(\mathbf{x})$: **توابع ویژه** (Eigenfunctions)

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

$$k(\mathbf{x},\mathbf{x}') = \sum_{i=1}^{\infty} \lambda_i \varphi_i(\mathbf{x}) \varphi_i(\mathbf{x}')$$

$\varphi_i(\mathbf{x})$: توابع ویژه (Eigenfunctions)

λ_i مقادیر ویژه (Eigenvalues)

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

$$k(\mathbf{x},\mathbf{x}') = \sum_{i=1}^{\infty} \lambda_i \varphi_i(\mathbf{x}) \varphi_i(\mathbf{x}')$$

$\varphi_i(\mathbf{x})$: **توابع ویژه** (Eigenfunctions)

λ_i **مقادیر ویژه** (Eigenvalues)

- **توجه: قضیه مرسر فقط می‌گوید که تابع انتخاب شده، کرنل ضرب داخلی است یا خیر.**

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

$$k(\mathbf{x},\mathbf{x}')=\sum_{i=1}^{\infty}\lambda_i\varphi_i(\mathbf{x})\varphi_i(\mathbf{x}')$$

$\varphi_i(\mathbf{x})$: توابع ویژه (Eigenfunctions)

λ_i مقادیر ویژه (Eigenvalues)

- توجه: قضیه مرسر فقط می‌گوید که تابع انتخاب شده، کرنل ضرب داخلی است یا خیر. این قضیه چیزی در مورد چگونگی تشکیل این تابع نمی‌گوید.

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

$$k(\mathbf{x},\mathbf{x}') = \sum_{i=1}^{\infty} \lambda_i \varphi_i(\mathbf{x}) \varphi_i(\mathbf{x}')$$

$\varphi_i(\mathbf{x})$: توابع ویژه (Eigenfunctions)

λ_i مقادیر ویژه (Eigenvalues)

- توجه: قضیه مرسر فقط می‌گوید که تابع انتخاب شده، کرنل ضرب داخلی است یا خیر. این قضیه چیزی در مورد چگونگی تشکیل این تابع نمی‌گوید.
- قضیه مرسر از آن جهت مهم است که تعداد کرنل‌ها را محدود می‌کند.

ماشین بردار پشتیبان (SVM)

۳- SVM با استفاده از توابع کرنل

$$k(\mathbf{x},\mathbf{x}')=\sum_{i=1}^{\infty}\lambda_i\varphi_i(\mathbf{x})\varphi_i(\mathbf{x}')$$

$\varphi_i(\mathbf{x})$: توابع ویژه (Eigenfunctions)

λ_i مقادیر ویژه (Eigenvalues)

- توجه: قضیه مرسر فقط می‌گوید که تابع انتخاب شده، کرنل ضرب داخلی است یا خیر. این قضیه چیزی در مورد چگونگی تشکیل این تابع نمی‌گوید.
- قضیه مرسر از آن جهت مهم است که تعداد کرنل‌ها را محدود می‌کند.

نتیجه گیری: بسط

$$k(\mathbf{x},\mathbf{x}_i)=\sum_{j=1}^{\infty}\varphi_j(\mathbf{x}_i)\,\varphi_j(\mathbf{x})$$

حالت خاصی از قضیه مرسر است با $\lambda_i=1$

ماشین بردار پشتیبان (SVM)

طراحی ماشین بردار پشتیبان:

ماشین بردار پشتیبان (SVM)

طراحی ماشین بردار پشتیبان:

– ابتدا توجه کنید که کرنل $k(\mathbf{x}, \mathbf{x}_i)$ به ما اجازه می‌دهد که سطح تصمیم‌گیری غیرخطی را در فضای ورودی تشکیل دهیم به‌طوری که شکل این سطح در فضای ویژگی‌ها، خطی است.

ماشین بردار پشتیبان (SVM)

طراحی ماشین بردار پشتیبان:

– ابتدا توجه کنید که کرنل $k(\mathbf{x}, \mathbf{x}_i)$ به ما اجازه می‌دهد که سطح تصمیم‌گیری غیرخطی را در فضای ورودی تشکیل دهیم به‌طوری که شکل این سطح در فضای ویژگی‌ها، خطی است.

– برای طراحی ماشین بردار پشتیبان فقط نیاز به استفاده از فرم دوگان بهینه سازی مقید داریم.

ماشین بردار پشتیبان (SVM)

طراحی ماشین بردار پشتیبان:

– ابتدا توجه کنید که کرنل $k(\mathbf{x},\mathbf{x}_i)$ به ما اجازه می‌دهد که سطح تصمیم‌گیری غیرخطی را در فضای ورودی تشکیل دهیم به‌طوری که شکل این سطح در فضای ویژگی‌ها، خطی است.

– برای طراحی ماشین بردار پشتیبان فقط نیاز به استفاده از فرم دوگان بهینه سازی مقید داریم.

با توجه به داده‌های آموزش $\{(\mathbf{x}_i, d_i)\}_{i=1}^{N}$ ، مطلوب است یافتن ضرایب لاگرانژ $\{\alpha_i\}_{i=1}^{N}$ که تابع هزینه زیر را بیشینه کند:

ماشین بردار پشتیبان (SVM)

طراحی ماشین بردار پشتیبان:

– ابتدا توجه کنید که کرنل $k(\mathbf{x},\mathbf{x}_i)$ به ما اجازه می‌دهد که سطح تصمیم‌گیری غیرخطی را در فضای ورودی تشکیل دهیم به‌طوری که شکل این سطح در فضای ویژگی‌ها، خطی است.

– برای طراحی ماشین بردار پشتیبان فقط نیاز به استفاده از فرم دوگان بهینه سازی مقید داریم.

با توجه به داده‌های آموزش $\{(\mathbf{x}_i, d_i)\}_{i=1}^{N}$ ، مطلوب است یافتن ضرایب لاگرانژ $\{\alpha_i\}_{i=1}^{N}$ که تابع هزینه زیر را بیشینه کند:

$$Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j k(\mathbf{x}_i, \mathbf{x}_j)$$

ماشین بردار پشتیبان (SVM)

طراحی ماشین بردار پشتیبان:

– ابتدا توجه کنید که کرنل $k(\mathbf{x},\mathbf{x}_i)$ به ما اجازه می‌دهد که سطح تصمیم‌گیری غیرخطی را در فضای ورودی تشکیل دهیم به‌طوری که شکل این سطح در فضای ویژگی‌ها، خطی است.

– برای طراحی ماشین بردار پشتیبان فقط نیاز به استفاده از فرم دوگان بهینه سازی مقید داریم.

با توجه به داده‌های آموزش $\{(\mathbf{x}_i, d_i)\}_{i=1}^{N}$ ، مطلوب است یافتن ضرایب لاگرانژ $\{\alpha_i\}_{i=1}^{N}$ که تابع هزینه زیر را بیشینه کند:

$$Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j k(\mathbf{x}_i, \mathbf{x}_j)$$

با درنظرگرفتن قیود

$$\sum_{i=1}^{N} \alpha_i d_i = 0$$

$$0 \leq \alpha_i \leq C \qquad i = 1,\ldots,N$$

ماشین بردار پشتیبان (SVM)

طراحی ماشین بردار پشتیبان:

– ابتدا توجه کنید که کرنل $k(\mathbf{x},\mathbf{x}_i)$ به ما اجازه می‌دهد که سطح تصمیم‌گیری غیرخطی را در فضای ورودی تشکیل دهیم به‌طوری که شکل این سطح در فضای ویژگی‌ها، خطی است.

– برای طراحی ماشین بردار پشتیبان فقط نیاز به استفاده از فرم دوگان بهینه سازی مقید داریم.

با توجه به داده‌های آموزش $\{(\mathbf{x}_i, d_i)\}_{i=1}^N$ ، مطلوب است یافتن ضرایب لاگرانژ $\{\alpha_i\}_{i=1}^N$ که تابع هزینه زیر را بیشینه کند:

$$Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2}\sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j d_i d_j k(\mathbf{x}_i, \mathbf{x}_j)$$

با درنظرگرفتن قیود

$$\sum_{i=1}^N \alpha_i d_i = 0$$

$$0 \le \alpha_i \le C \qquad i = 1,\dots,N$$

که در آن $C > 0$ توسط طراح تعیین می‌شود.

ماشین بردار پشتیبان (SVM)

طراحی ماشین بردار پشتیبان:

– ابتدا توجه کنید که کرنل $k(\mathbf{x}, \mathbf{x}_i)$ به ما اجازه می‌دهد که سطح تصمیم‌گیری غیرخطی را در فضای ورودی تشکیل دهیم به‌طوری که شکل این سطح در فضای ویژگی‌ها، خطی است.

– برای طراحی ماشین بردار پشتیبان فقط نیاز به استفاده از فرم دوگان بهینه سازی مقید داریم.

با توجه به داده‌های آموزش $\{(\mathbf{x}_i, d_i)\}_{i=1}^{N}$ ، مطلوب است یافتن ضرایب لاگرانژ $\{\alpha_i\}_{i=1}^{N}$ که تابع هزینه زیر را بیشینه کند:

$$Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j k(\mathbf{x}_i, \mathbf{x}_j)$$

با درنظرگرفتن قیود

$$\sum_{i=1}^{N} \alpha_i d_i = 0$$

$$0 \le \alpha_i \le C \qquad i = 1, \dots, N$$

که در آن $C > 0$ توسط طراح تعیین می‌شود.

– توجه کنید که این مساله دوگان دقیقا مثل حالت «الگوهای جداناپذیر» است با این تفاوت که به‌جای ضرب داخلی $\mathbf{x}_i^T \mathbf{x}_j$ از کرنل $k(\mathbf{x}, \mathbf{x}_i)$ استفاده شده است.

ماشین بردار پشتیبان (SVM)

- مثال‌هایی از ماشین بردار پشتیبان:

ماشین بردار پشتیبان (SVM)

– مثال‌هایی از ماشین بردار پشتیبان:

Type of support vector machine	Mercer kernel $k(\mathbf{x}, \mathbf{x}_i), i = 1, 2, ..., N$	Comments
Polynomial learning machine	$(\mathbf{x}^T\mathbf{x}_i + 1)^p$	Power p is specified *a priori* by the user
Radial-basis-function network	$\exp\left(-\frac{1}{2\sigma^2}\lVert\mathbf{x} - \mathbf{x}_i\rVert^2\right)$	The width σ^2, common to all the kernels, is specified *a priori* by the user
Two-layer perceptron	$\tanh(\beta_0\mathbf{x}^T\mathbf{x}_i + \beta_1)$	Mercer's theorem is satisfied only for some values of β_0 and β_1

ماشین بردار پشتیبان (SVM)

- مثال‌هایی از ماشین بردار پشتیبان:

Type of support vector machine	Mercer kernel $k(\mathbf{x}, \mathbf{x}_i), i = 1, 2, ..., N$	Comments
Polynomial learning machine	$(\mathbf{x}^T\mathbf{x}_i + 1)^p$	Power p is specified *a priori* by the user
Radial-basis-function network	$\exp\left(-\frac{1}{2\sigma^2}\|\mathbf{x} - \mathbf{x}_i\|^2\right)$	The width σ^2, common to all the kernels, is specified *a priori* by the user
Two-layer perceptron	$\tanh(\beta_0\mathbf{x}^T\mathbf{x}_i + \beta_1)$	Mercer's theorem is satisfied only for some values of β_0 and β_1

۱- کرنل‌های مرسر چندجمله ای و RBF همواره قضیه مرسر را برآورده می‌کند.

ماشین بردار پشتیبان (SVM)

- مثال‌هایی از ماشین بردار پشتیبان:

Type of support vector machine	Mercer kernel $k(\mathbf{x}, \mathbf{x}_i), i = 1, 2, ..., N$	Comments
Polynomial learning machine	$(\mathbf{x}^T\mathbf{x}_i + 1)^p$	Power p is specified *a priori* by the user
Radial-basis-function network	$\exp\left(-\frac{1}{2\sigma^2}\|\mathbf{x} - \mathbf{x}_i\|^2\right)$	The width σ^2, common to all the kernels, is specified *a priori* by the user
Two-layer perceptron	$\tanh(\beta_0\mathbf{x}^T\mathbf{x}_i + \beta_1)$	Mercer's theorem is satisfied only for some values of β_0 and β_1

۱- کرنل‌های مرسر چندجمله ای و RBF همواره قضیه مرسر را برآورده می‌کند.

۲- کرنل مرسر tanh به این واقعیت اشاره دارد که تعیین کرنل مرسر همواره کار ساده‌ای نیست.

ماشین بردار پشتیبان (SVM)

- مثال‌هایی از ماشین بردار پشتیبان:

Type of support vector machine	Mercer kernel $k(\mathbf{x}, \mathbf{x}_i),\ i = 1, 2, \ldots, N$	Comments
Polynomial learning machine	$(\mathbf{x}^T\mathbf{x}_i + 1)^p$	Power p is specified *a priori* by the user
Radial-basis-function network	$\exp\left(-\frac{1}{2\sigma^2}\lVert\mathbf{x} - \mathbf{x}_i\rVert^2\right)$	The width σ^2, common to all the kernels, is specified *a priori* by the user
Two-layer perceptron	$\tanh(\beta_0\mathbf{x}^T\mathbf{x}_i + \beta_1)$	Mercer's theorem is satisfied only for some values of β_0 and β_1

۱- کرنل‌های مرسر چندجمله ای و RBF همواره قضیه مرسر را برآورده می‌کنند.

۲- کرنل مرسر tanh به این واقعیت اشاره دارد که تعیین کرنل مرسر همواره کار ساده‌ای نیست.

۳- برای هر سه نوع ماشین، ابعاد فضای ویژگی‌ها (تعداد سلول‌ها) توسط تعداد بردارهای پشتیبان که از حل مساله بهینه‌سازی مقید به‌دست می‌آید، تعیین می‌شود.

ماشین بردار پشتیبان (SVM)

- مثال‌هایی از ماشین بردار پشتیبان:

Type of support vector machine	Mercer kernel $k(\mathbf{x}, \mathbf{x}_i), i = 1, 2, ..., N$	Comments
Polynomial learning machine	$(\mathbf{x}^T\mathbf{x}_i + 1)^p$	Power p is specified *a priori* by the user
Radial-basis-function network	$\exp\left(-\frac{1}{2\sigma^2}\lVert\mathbf{x} - \mathbf{x}_i\rVert^2\right)$	The width σ^2, common to all the kernels, is specified *a priori* by the user
Two-layer perceptron	$\tanh(\beta_0\mathbf{x}^T\mathbf{x}_i + \beta_1)$	Mercer's theorem is satisfied only for some values of β_0 and β_1

۱- کرنل‌های مرسر چندجمله ای و RBF همواره قضیه مرسر را برآورده می‌کند.

۲- کرنل مرسر tanh به این واقعیت اشاره دارد که تعیین کرنل مرسر همواره کار ساده‌ای نیست.

۳- برای هر سه نوع ماشین، ابعاد فضای ویژگی‌ها (تعداد سلول‌ها) توسط تعداد بردارهای پشتیبان که از حل مساله بهینه‌سازی مقید به‌دست می‌آید، تعیین می‌شود.

۴- در ماشین بردار پشتیبان، برخلاف RBF و MLP، نیازی به استفاده از روش‌های ابتکاری نیست.

ماشین بردار پشتیبان (SVM)

- مثال‌هایی از ماشین بردار پشتیبان:

Type of support vector machine	Mercer kernel $k(\mathbf{x}, \mathbf{x}_i), i = 1, 2, ..., N$	Comments
Polynomial learning machine	$(\mathbf{x}^T\mathbf{x}_i + 1)^p$	Power p is specified *a priori* by the user
Radial-basis-function network	$\exp\left(-\frac{1}{2\sigma^2}\lVert\mathbf{x} - \mathbf{x}_i\rVert^2\right)$	The width σ^2, common to all the kernels, is specified *a priori* by the user
Two-layer perceptron	$\tanh(\beta_0\mathbf{x}^T\mathbf{x}_i + \beta_1)$	Mercer's theorem is satisfied only for some values of β_0 and β_1

۱- کرنل‌های مرسر چندجمله ای و RBF همواره قضیه مرسر را برآورده می‌کند.

۲- کرنل مرسر tanh به این واقعیت اشاره دارد که تعیین کرنل مرسر همواره کار ساده‌ای نیست.

۳- برای هر سه نوع ماشین، ابعاد فضای ویژگی‌ها (تعداد سلول‌ها) توسط تعداد بردارهای پشتیبان که از حل مساله بهینه‌سازی مقید به‌دست می‌آید، تعیین می‌شود.

۴- در ماشین بردار پشتیبان، برخلاف RBF و MLP، نیازی به استفاده از روش‌های ابتکاری نیست.

۵- در ماشین بردار پشتیبانRBF، تعداد RBF ها و مرکز آن‌ها به‌ترتیب توسط بردارهای پشتیبان و مقدار آن‌ها تعیین می‌شود.

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

x_1, x_2	d
$(-1,-1)$	-1
$(-1,+1)$	$+1$
$(+1,-1)$	$+1$
$(+1,+1)$	-1

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

x_1, x_2	d
$(-1,-1)$	-1
$(-1,+1)$	$+1$
$(+1,-1)$	$+1$
$(+1,+1)$	-1

-کرنل درنظرگرفته شده

$$k(\mathbf{x},\mathbf{x}_i) = (1 + \mathbf{x}^T\mathbf{x}_i)^2$$

$$\mathbf{x} = [x_1 \quad x_2]^T \quad \mathbf{x}_i = [x_{i1} \quad x_{i2}]^T$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

x_1, x_2	d
(−1,−1)	−1
(−1,+1)	+1
(+1, −1)	+1
(+1,+1)	−1

-کرنل درنظرگرفته شده

$$k(\mathbf{x}, \mathbf{x}_i) = (1 + \mathbf{x}^T \mathbf{x}_i)^2$$

$$\mathbf{x} = [x_1 \quad x_2]^T \quad \mathbf{x}_i = [x_{i1} \quad x_{i2}]^T$$

بنابراین

$$k(\mathbf{x}, \mathbf{x}_i) = 1 + x_1^2 x_{i1}^2 + 2x_1 x_2 x_{i1} x_{i2} + x_2^2 x_{i2}^2 + 2x_1 x_{i1} + 2x_2 x_{i2}$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

x_1, x_2	d
(−1,−1)	−1
(−1,+1)	+1
(+1, −1)	+1
(+1,+1)	−1

-کرنل درنظرگرفته شده

$$k(\mathbf{x},\mathbf{x}_i) = (1+\mathbf{x}^T\mathbf{x}_i)^2$$

$$\mathbf{x} = [x_1 \quad x_2]^T \quad \mathbf{x}_i = [x_{i1} \quad x_{i2}]^T$$

بنابراین

$$k(\mathbf{x},\mathbf{x}_i) = 1 + x_1^2x_{i1}^2 + 2x_1x_2x_{i1}x_{i2} + x_2^2x_{i2}^2 + 2x_1x_{i1} + 2x_2x_{i2}$$

در نتیجه می‌توان استنباط‌کرد که تصویر انتقال یافته $\mathbf{x}$ به فضای ویژگی‌ها برابر است با

$$\boldsymbol{\phi}(\mathbf{x}) = [1, x_1^2, \sqrt{2}x_1x_2, x_2^2, \sqrt{2}x_1, \sqrt{2}x_2]^T$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

x_1, x_2	d
(−1,−1)	−1
(−1,+1)	+1
(+1, −1)	+1
(+1,+1)	−1

-کرنل درنظرگرفته شده

$$k(\mathbf{x}, \mathbf{x}_i) = (1 + \mathbf{x}^T \mathbf{x}_i)^2$$

$$\mathbf{x} = [x_1 \quad x_2]^T \quad \mathbf{x}_i = [x_{i1} \quad x_{i2}]^T$$

بنابراین

$$k(\mathbf{x}, \mathbf{x}_i) = 1 + x_1^2 x_{i1}^2 + 2x_1 x_2 x_{i1} x_{i2} + x_2^2 x_{i2}^2 + 2x_1 x_{i1} + 2x_2 x_{i2}$$

در نتیجه می‌توان استنباط‌کرد که تصویر انتقال یافته x به فضای ویژگی‌ها برابر است با

$$\boldsymbol{\phi}(\mathbf{x}) = [1, x_1^2, \sqrt{2}x_1 x_2, x_2^2, \sqrt{2}x_1, \sqrt{2}x_2]^T$$

و به همین ترتیب برای $\mathbf{x}_i$

$$\boldsymbol{\phi}(\mathbf{x}_i) = [1, x_{i1}^2, \sqrt{2}x_{i1} x_{i2}, x_{i2}^2, \sqrt{2}x_{i1}, \sqrt{2}x_{i2}]^T, \qquad i = 1, 2, 3, 4$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

x_1, x_2	d
(−1,−1)	−1
(−1,+1)	+1
(+1, −1)	+1
(+1,+1)	−1

-کرنل درنظرگرفته شده

$$k(\mathbf{x},\mathbf{x}_i) = (1+\mathbf{x}^T\mathbf{x}_i)^2$$

$$\mathbf{x} = [x_1 \quad x_2]^T \quad \mathbf{x}_i = [x_{i1} \quad x_{i2}]^T$$

بنابراین

$$k(\mathbf{x}, \mathbf{x}_i) = 1 + x_1^2x_{i1}^2 + 2x_1x_2x_{i1}x_{i2} + x_2^2x_{i2}^2 + 2x_1x_{i1} + 2x_2x_{i2}$$

در نتیجه می‌توان استنباط‌کرد که تصویر انتقال یافته x به فضای ویژگی‌ها برابر است با

$$\boldsymbol{\phi}(\mathbf{x}) = [1, x_1^2, \sqrt{2}x_1x_2, x_2^2, \sqrt{2}x_1, \sqrt{2}x_2]^T$$

و به همین ترتیب برای $\mathbf{x}_i$

$$\boldsymbol{\phi}(\mathbf{x}_i) = [1, x_{i1}^2, \sqrt{2}x_{i1}x_{i2}, x_{i2}^2, \sqrt{2}x_{i1}, \sqrt{2}x_{i2}]^T, \qquad i = 1, 2, 3, 4$$

ماتریس کرنل (گرام)

$$\mathbf{K} = \{k(\mathbf{x}_i, \mathbf{x}_j)\}_{i,j=1}^N$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

x_1, x_2	d
(−1,−1)	−1
(−1,+1)	+1
(+1, −1)	+1
(+1,+1)	−1

-کرنل درنظرگرفته شده

$$k(\mathbf{x}, \mathbf{x}_i) = (1 + \mathbf{x}^T \mathbf{x}_i)^2$$

$$\mathbf{x} = [x_1 \quad x_2]^T \quad \mathbf{x}_i = [x_{i1} \quad x_{i2}]^T$$

بنابراین

$$k(\mathbf{x}, \mathbf{x}_i) = 1 + x_1^2 x_{i1}^2 + 2x_1 x_2 x_{i1} x_{i2} + x_2^2 x_{i2}^2 + 2x_1 x_{i1} + 2x_2 x_{i2}$$

در نتیجه می‌توان استنباط‌کرد که تصویر انتقال یافته $\mathbf{x}$ به فضای ویژگی‌ها برابر است با

$$\boldsymbol{\phi}(\mathbf{x}) = [1, x_1^2, \sqrt{2}x_1 x_2, x_2^2, \sqrt{2}x_1, \sqrt{2}x_2]^T$$

و به همین ترتیب برای $\mathbf{x}_i$

$$\boldsymbol{\phi}(\mathbf{x}_i) = [1, x_{i1}^2, \sqrt{2}x_{i1} x_{i2}, x_{i2}^2, \sqrt{2}x_{i1}, \sqrt{2}x_{i2}]^T, \qquad i = 1, 2, 3, 4$$

ماتریس کرنل (گرام)

$$\mathbf{K} = \{k(\mathbf{x}_i, \mathbf{x}_j)\}_{i,j=1}^{N} \Rightarrow \mathbf{K} = \begin{bmatrix} 9 & 1 & 1 & 1 \\ 1 & 9 & 1 & 1 \\ 1 & 1 & 9 & 1 \\ 1 & 1 & 1 & 9 \end{bmatrix}$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

تابع هزینه مساله دوگان بهینه‌سازی

$$Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j k(\mathbf{x}_i, \mathbf{x}_j)$$

$$\sum_{i=1}^{N} \alpha_i d_i = 0$$

$$0 \leq \alpha_i \leq C \qquad i = 1, \ldots, N$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

تابع هزینه مساله دوگان بهینه‌سازی

$$Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j k(\mathbf{x}_i, \mathbf{x}_j)$$

$$\sum_{i=1}^{N} \alpha_i d_i = 0$$

$$0 \le \alpha_i \le C \qquad i = 1, \dots, N$$

$$Q(\alpha) = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2}(9\alpha_1^2 - 2\alpha_1\alpha_2 - 2\alpha_1\alpha_3 + 2\alpha_1\alpha_4$$
$$+ 9\alpha_2^2 + 2\alpha_2\alpha_3 - 2\alpha_2\alpha_4 + 9\alpha_3^2 - 2\alpha_3\alpha_4 + 9\alpha_4^2)$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

تابع هزینه مساله دوگان بهینه‌سازی

$$Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j k(\mathbf{x}_i, \mathbf{x}_j)$$

$$\sum_{i=1}^{N} \alpha_i d_i = 0$$

$$0 \le \alpha_i \le C \qquad i = 1, \dots, N$$

$$Q(\alpha) = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2}(9\alpha_1^2 - 2\alpha_1\alpha_2 - 2\alpha_1\alpha_3 + 2\alpha_1\alpha_4 + 9\alpha_2^2 + 2\alpha_2\alpha_3 - 2\alpha_2\alpha_4 + 9\alpha_3^2 - 2\alpha_3\alpha_4 + 9\alpha_4^2)$$

بهینه‌سازی این تابع هزینه برحسب ضرایب لاگرانژ، نتیجه می‌دهد:

$$\begin{aligned}
9\alpha_1 - \alpha_2 - \alpha_3 + \alpha_4 &= 1 \\
-\alpha_1 + 9\alpha_2 + \alpha_3 - \alpha_4 &= 1 \\
-\alpha_1 + \alpha_2 + 9\alpha_3 - \alpha_4 &= 1 \\
\alpha_1 - \alpha_2 - \alpha_3 + 9\alpha_4 &= 1
\end{aligned}$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

بهینه‌سازی این تابع هزینه برحسب ضرایب لاگرانژ، نتیجه می‌دهد:

$$
\begin{aligned}
9\alpha_1 - \alpha_2 - \alpha_3 + \alpha_4 &= 1 \\
-\alpha_1 + 9\alpha_2 + \alpha_3 - \alpha_4 &= 1 \\
-\alpha_1 + \alpha_2 + 9\alpha_3 - \alpha_4 &= 1 \\
\alpha_1 - \alpha_2 - \alpha_3 + 9\alpha_4 &= 1
\end{aligned}
$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

بهینه‌سازی این تابع هزینه برحسب ضرایب لاگرانژ، نتیجه می‌دهد:

$$\begin{aligned}
9\alpha_1 - \alpha_2 - \alpha_3 + \alpha_4 &= 1 \\
-\alpha_1 + 9\alpha_2 + \alpha_3 - \alpha_4 &= 1 \\
-\alpha_1 + \alpha_2 + 9\alpha_3 - \alpha_4 &= 1 \\
\alpha_1 - \alpha_2 - \alpha_3 + 9\alpha_4 &= 1
\end{aligned}$$

بنابراین، مقدار بهینه این ضرایب برابراند با

$$\alpha_{o,1} = \alpha_{o,2} = \alpha_{o,3} = \alpha_{o,4} = \frac{1}{8}$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

بهینه‌سازی این تابع هزینه برحسب ضرایب لاگرانژ، نتیجه می‌دهد:

$$\begin{aligned} 9\alpha_1 - \alpha_2 - \alpha_3 + \alpha_4 &= 1 \\ -\alpha_1 + 9\alpha_2 + \alpha_3 - \alpha_4 &= 1 \\ -\alpha_1 + \alpha_2 + 9\alpha_3 - \alpha_4 &= 1 \\ \alpha_1 - \alpha_2 - \alpha_3 + 9\alpha_4 &= 1 \end{aligned}$$

بنابراین، مقدار بهینه این ضرایب برابراند با

$$\alpha_{o,1} = \alpha_{o,2} = \alpha_{o,3} = \alpha_{o,4} = \frac{1}{8}$$

یعنی این که هر چهار بردار ورودی، بردار پشتیبان هستند. یعنی رابطه

$$\sum_{i=1}^{N} \alpha_i d_i = 0$$

برای تمامی داده‌ها، برقرار است.

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

مقدار بهینه تابع هزینه

$$Q_o(\alpha) = \frac{1}{4}$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

مقدار بهینه تابع هزینه

$$Q_o(\alpha) = \frac{1}{4}$$

از آن جا می‌توان نتیجه گرفت که

$$\frac{1}{2}\|\mathbf{w}_o\|^2 = \frac{1}{4} \quad \Rightarrow \quad \|\mathbf{w}_o\| = \frac{1}{\sqrt{2}}$$

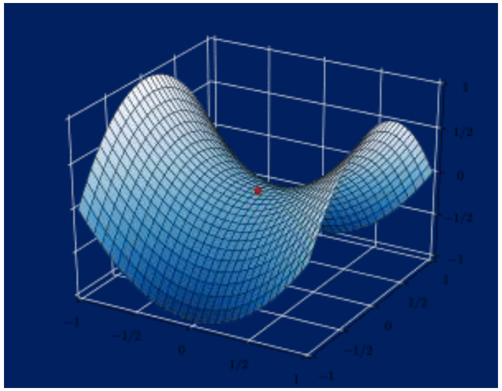
ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

مقدار بهینه تابع هزینه

$$Q_o(\alpha) = \frac{1}{4}$$

از آن جا می‌توان نتیجه گرفت که



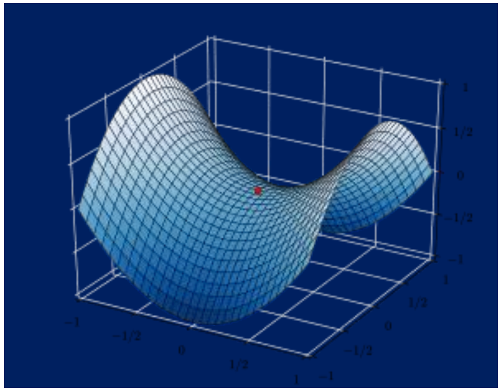
$$\frac{1}{2}\|\mathbf{w}_o\|^2 = \frac{1}{4} \quad \Rightarrow \quad \|\mathbf{w}_o\| = \frac{1}{\sqrt{2}}$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

مقدار بهینه تابع هزینه

$$Q_o(\alpha) = \frac{1}{4}$$



از آن جا می‌توان نتیجه گرفت که

$$\frac{1}{2}\|\mathbf{w}_o\|^2 = \frac{1}{4} \quad \Rightarrow \quad \|\mathbf{w}_o\| = \frac{1}{\sqrt{2}}$$

با استفاده از

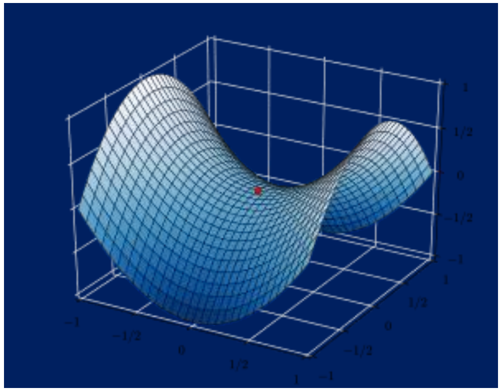
$$\mathbf{w} = \sum_{i=1}^{Ns} \alpha_i d_i \Phi(\mathbf{x}_i)$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

مقدار بهینه تابع هزینه

$$Q_o(\alpha) = \frac{1}{4}$$



از آن جا می‌توان نتیجه گرفت که

$$\frac{1}{2}\|\mathbf{w}_o\|^2 = \frac{1}{4} \quad \Rightarrow \quad \|\mathbf{w}_o\| = \frac{1}{\sqrt{2}}$$

با استفاده از

$$\mathbf{w} = \sum_{i=1}^{Ns} \alpha_i d_i \Phi(\mathbf{x}_i)$$

خواهیم داشت

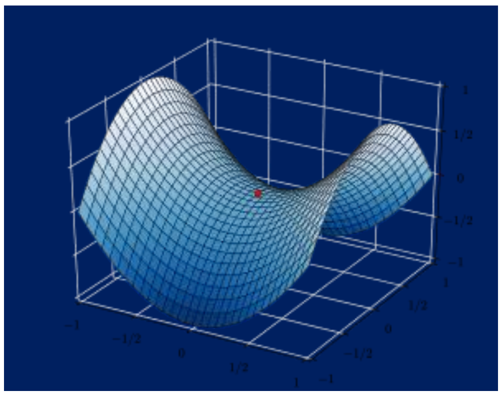
$$\mathbf{w}_o = \frac{1}{8}[-\varphi(\mathbf{x}_1) + \varphi(\mathbf{x}_2) + \varphi(\mathbf{x}_3) - \varphi(\mathbf{x}_4)]$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

مقدار بهینه تابع هزینه

$$Q_o(\alpha) = \frac{1}{4}$$



از آن جا می‌توان نتیجه گرفت که

$$\frac{1}{2}\|\mathbf{w}_o\|^2 = \frac{1}{4} \quad \Rightarrow \quad \|\mathbf{w}_o\| = \frac{1}{\sqrt{2}}$$

با استفاده از

$$\mathbf{w} = \sum_{i=1}^{Ns} \alpha_i d_i \Phi(\mathbf{x}_i)$$

خواهیم داشت

$$\mathbf{w}_o = \frac{1}{8}[-\varphi(\mathbf{x}_1) + \varphi(\mathbf{x}_2) + \varphi(\mathbf{x}_3) - \varphi(\mathbf{x}_4)]$$

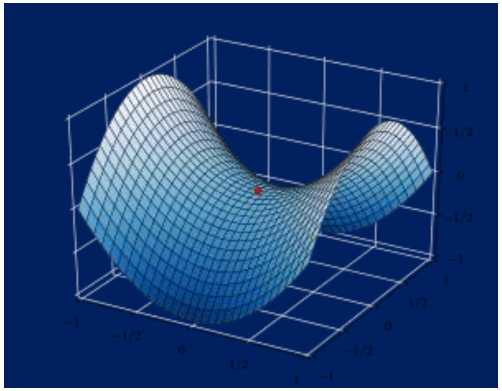
$$= \frac{1}{8}\left[-\begin{bmatrix} 1 \\ 1 \\ \sqrt{2} \\ 1 \\ -\sqrt{2} \\ -\sqrt{2} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -\sqrt{2} \\ 1 \\ -\sqrt{2} \\ \sqrt{2} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -\sqrt{2} \\ 1 \\ \sqrt{2} \\ -\sqrt{2} \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ \sqrt{2} \\ 1 \\ \sqrt{2} \\ \sqrt{2} \end{bmatrix} \right]$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

مقدار بهینه تابع هزینه

$$Q_o(\alpha) = \frac{1}{4}$$



از آن جا می‌توان نتیجه گرفت که

$$\frac{1}{2}\|\mathbf{w}_o\|^2 = \frac{1}{4} \quad \Rightarrow \quad \|\mathbf{w}_o\| = \frac{1}{\sqrt{2}}$$

با استفاده از

$$\mathbf{w} = \sum_{i=1}^{Ns} \alpha_i d_i \Phi(\mathbf{x}_i)$$

خواهیم داشت

$$\mathbf{w}_o = \frac{1}{8}[-\varphi(\mathbf{x}_1) + \varphi(\mathbf{x}_2) + \varphi(\mathbf{x}_3) - \varphi(\mathbf{x}_4)]$$

$$= \frac{1}{8}\left[-\begin{bmatrix} 1 \\ 1 \\ \sqrt{2} \\ 1 \\ -\sqrt{2} \\ -\sqrt{2} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -\sqrt{2} \\ 1 \\ -\sqrt{2} \\ \sqrt{2} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -\sqrt{2} \\ 1 \\ \sqrt{2} \\ -\sqrt{2} \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ \sqrt{2} \\ 1 \\ \sqrt{2} \\ \sqrt{2} \end{bmatrix}\right] = \begin{bmatrix} 0 \\ 0 \\ -1/\sqrt{2} \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

ابرصفحه بهینه

$$\mathbf{w}_o^T \boldsymbol{\phi}(\mathbf{x}) = 0$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

ابرصفحه بهینه

$$\mathbf{w}_o^T \boldsymbol{\phi}(\mathbf{x}) = 0 \quad \Rightarrow \left[0, 0, \frac{-1}{\sqrt{2}}, 0, 0, 0\right] \begin{bmatrix} 1 \\ x_1^2 \\ \sqrt{2}x_1x_2 \\ x_2^2 \\ \sqrt{2}x_1 \\ \sqrt{2}x_2 \end{bmatrix} = 0$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

ابرصفحه بهینه

$$\mathbf{w}_o^T \boldsymbol{\phi}(\mathbf{x}) = 0 \quad \Rightarrow \quad \left[0, 0, \frac{-1}{\sqrt{2}}, 0, 0, 0\right] \begin{bmatrix} 1 \\ x_1^2 \\ \sqrt{2}x_1x_2 \\ x_2^2 \\ \sqrt{2}x_1 \\ \sqrt{2}x_2 \end{bmatrix} = 0 \quad \Rightarrow \quad -x_1x_2 = 0$$

ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

ابرصفحه بهینه

$$\mathbf{w}_o^T\boldsymbol{\phi}(\mathbf{x}) = 0 \quad \Rightarrow \quad \left[0, 0, \frac{-1}{\sqrt{2}}, 0, 0, 0\right]\begin{bmatrix} 1 \\ x_1^2 \\ \sqrt{2}x_1x_2 \\ x_2^2 \\ \sqrt{2}x_1 \\ \sqrt{2}x_2 \end{bmatrix} = 0 \quad \Rightarrow \quad -x_1x_2 = 0$$

x_1, x_2	y
(−1,−1)	−1
(−1,+1)	+1
(+1, −1)	+1
(+1,+1)	−1

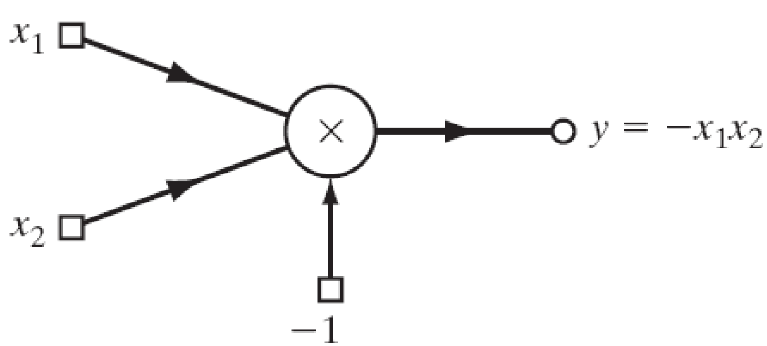
ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

ابرصفحه بهینه

$$\mathbf{w}_o^T \boldsymbol{\phi}(\mathbf{x}) = 0 \Rightarrow \left[0, 0, \frac{-1}{\sqrt{2}}, 0, 0, 0\right] \begin{bmatrix} 1 \\ x_1^2 \\ \sqrt{2}x_1x_2 \\ x_2^2 \\ \sqrt{2}x_1 \\ \sqrt{2}x_2 \end{bmatrix} = 0 \Rightarrow -x_1x_2 = 0$$

x_1, x_2	y
(–1,–1)	–1
(–1,+1)	+1
(+1, –1)	+1
(+1,+1)	–1



ماشین چندجمله‌ای

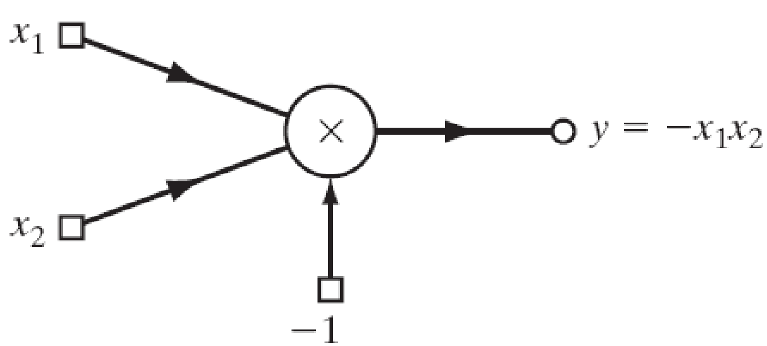
ماشین بردار پشتیبان (SVM)

مثال: مساله XOR

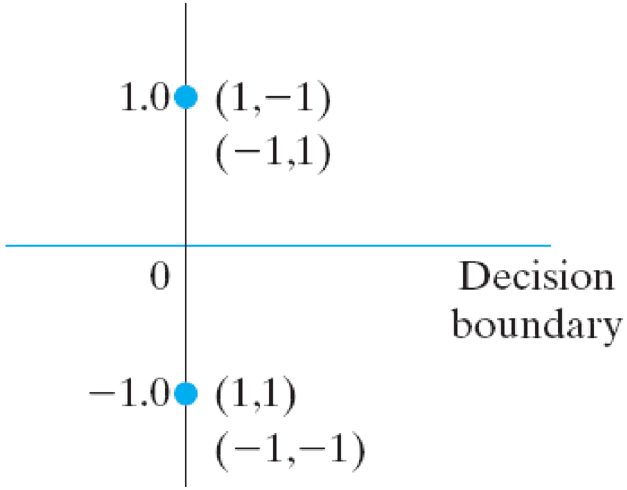
ابرصفحه بهینه

$$\mathbf{w}_o^T\boldsymbol{\phi}(\mathbf{x}) = 0 \;\Rightarrow\; \left[0, 0, \frac{-1}{\sqrt{2}}, 0, 0, 0\right]\begin{bmatrix}1\\ x_1^2\\ \sqrt{2}x_1x_2\\ x_2^2\\ \sqrt{2}x_1\\ \sqrt{2}x_2\end{bmatrix} = 0 \;\Rightarrow\; -x_1x_2 = 0$$

x_1, x_2	y
(−1,−1)	−1
(−1,+1)	+1
(+1, −1)	+1
(+1,+1)	−1



ماشین چندجمله‌ای



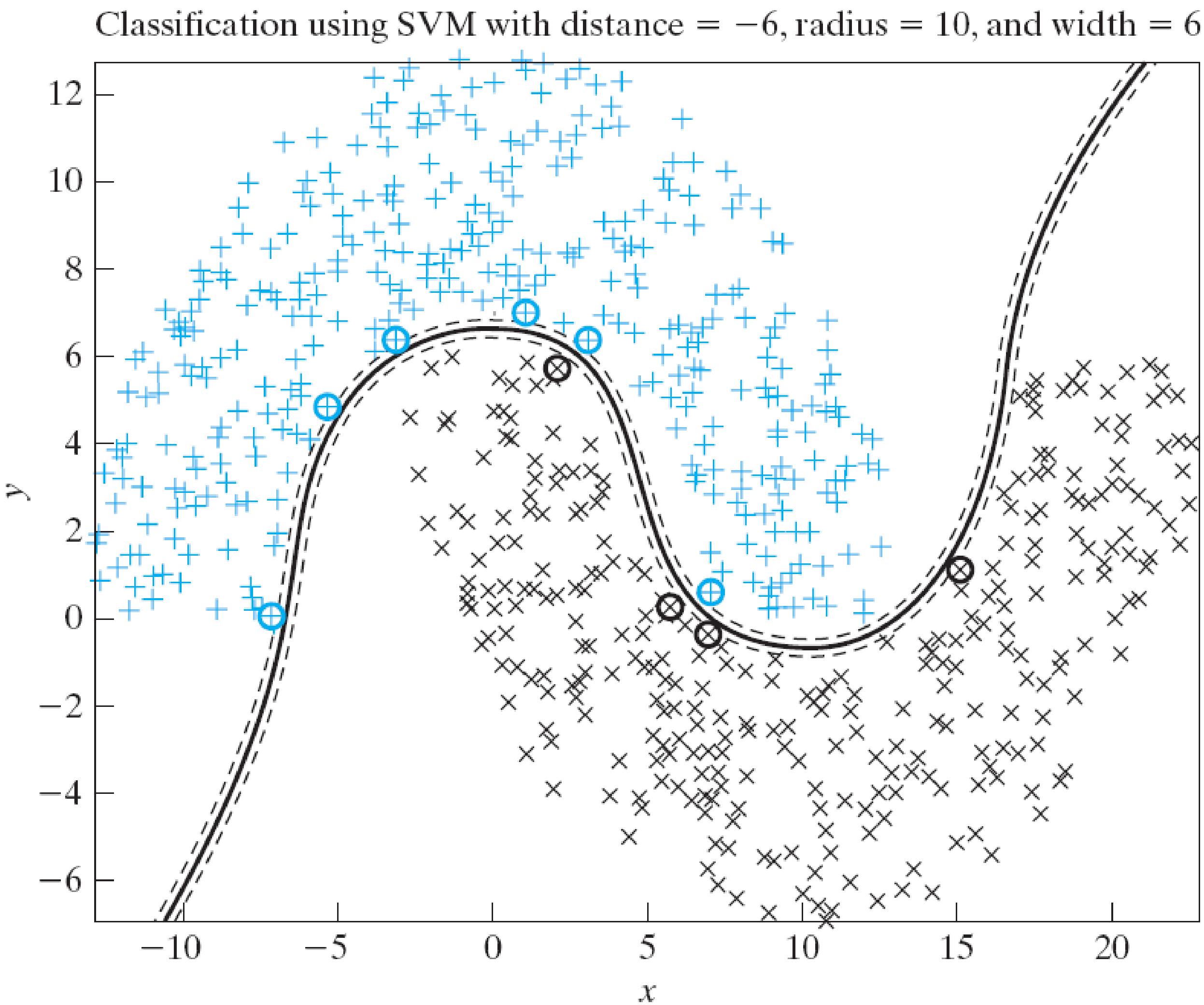
تصویر القاشده به فضای ویژگی‌ها

ماشین بردار پشتیبان (SVM)

مثال: کلاسه‌بندی کلاس‌های ماه شکل

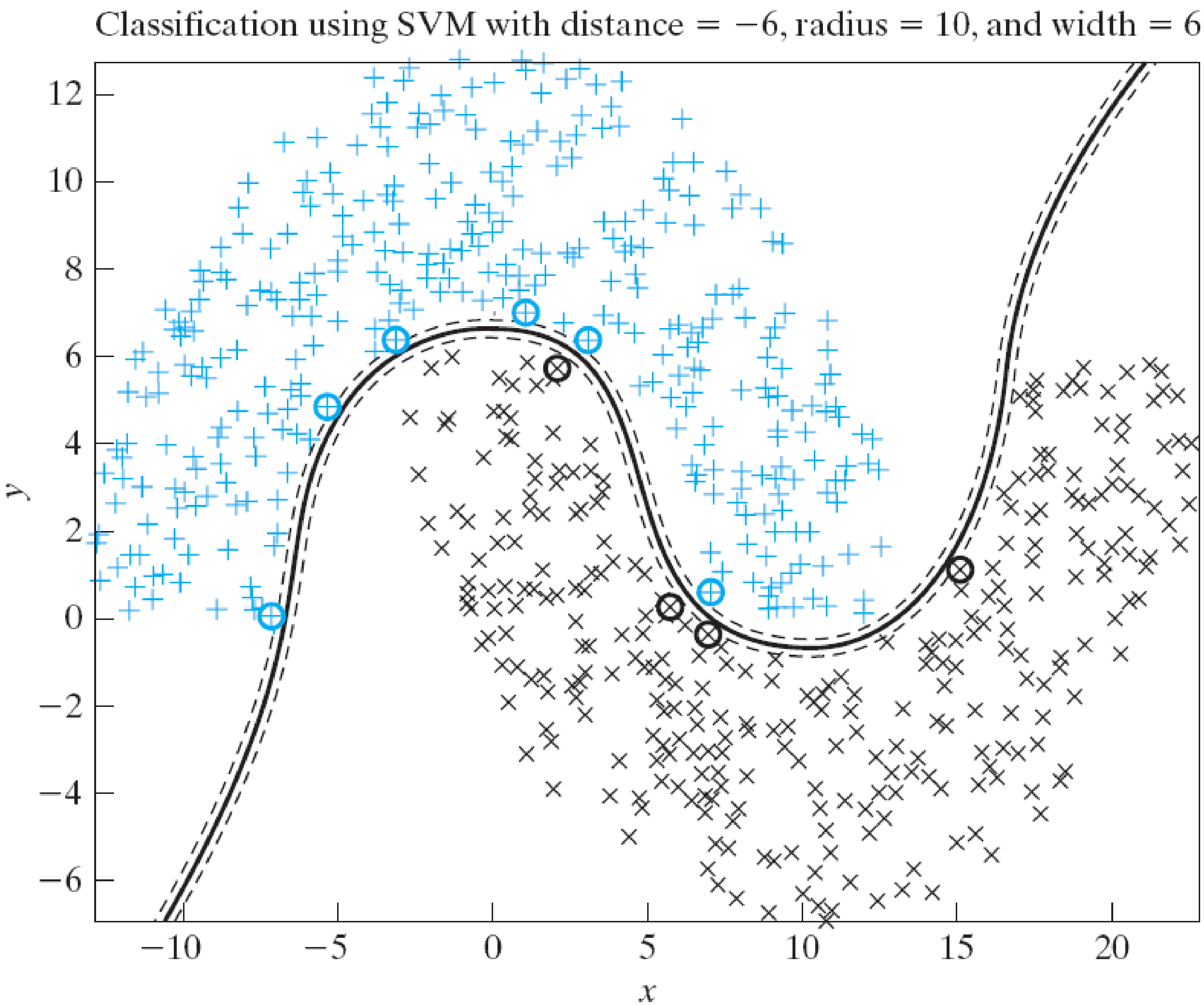
ماشین بردار پشتیبان (SVM)

مثال: کلاسه‌بندی کلاس‌های ماه شکل



ماشین بردار پشتیبان (SVM)

مثال: کلاسه‌بندی کلاس‌های ماه شکل

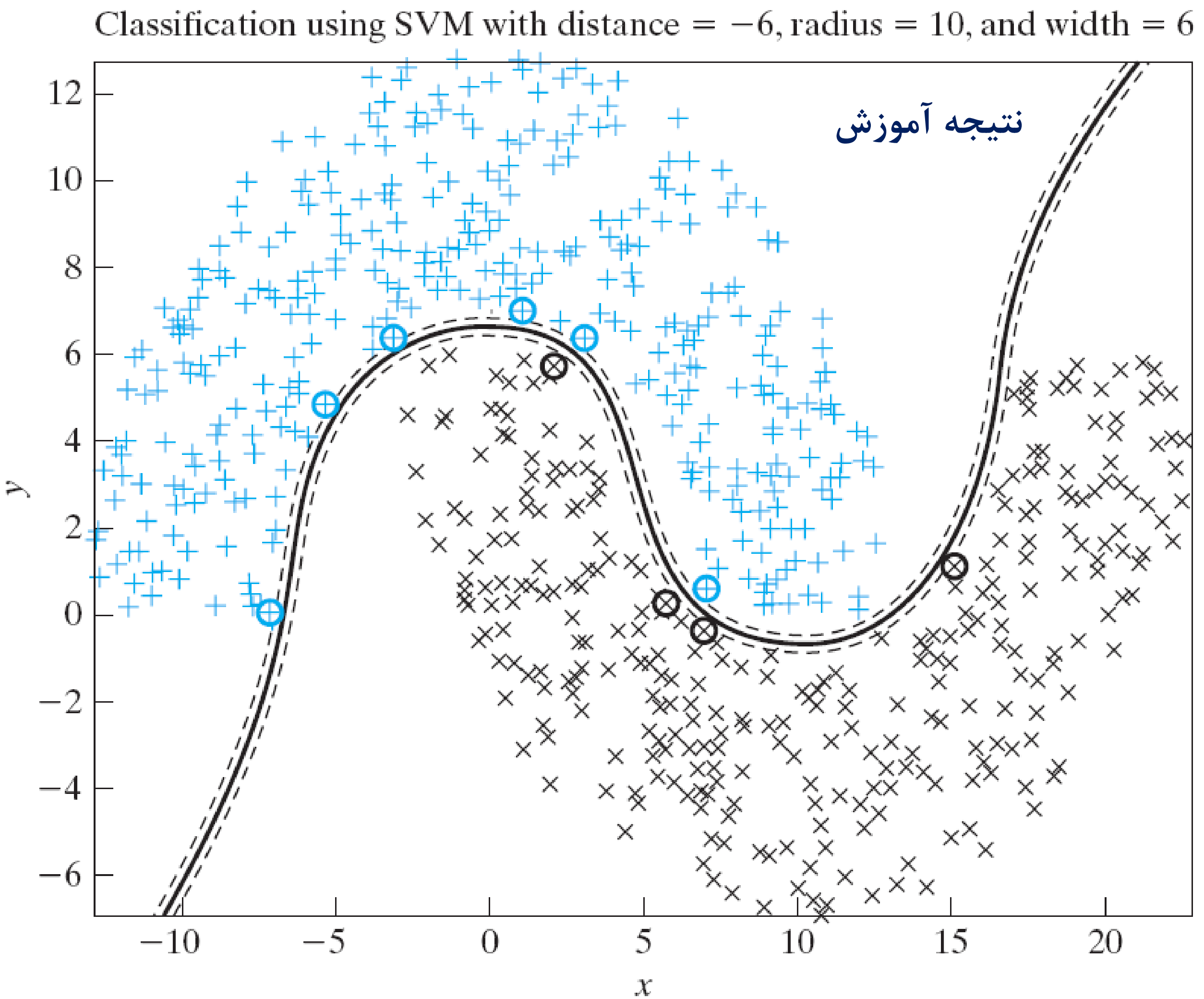


۳۰۰ داده برای آموزش
۲۰۰۰ داده برای آزمایش

$C = \infty$

ماشین بردار پشتیبان (SVM)

مثال: کلاسه‌بندی کلاس‌های ماه شکل

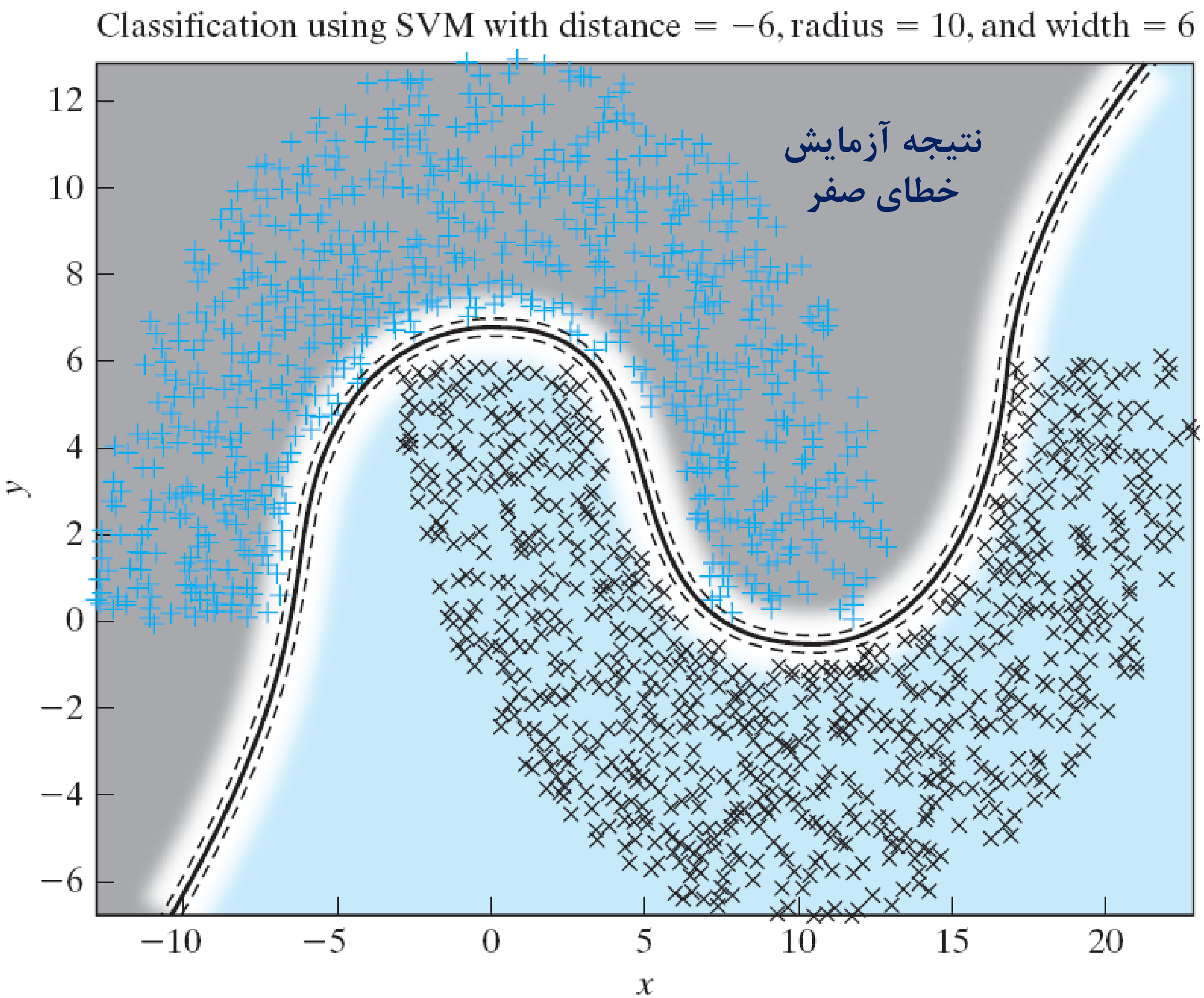


۳۰۰ داده برای آموزش
۲۰۰۰ داده برای آزمایش

$C = \infty$

ماشین بردار پشتیبان (SVM)

مثال: کلاسه‌بندی کلاس‌های ماه شکل

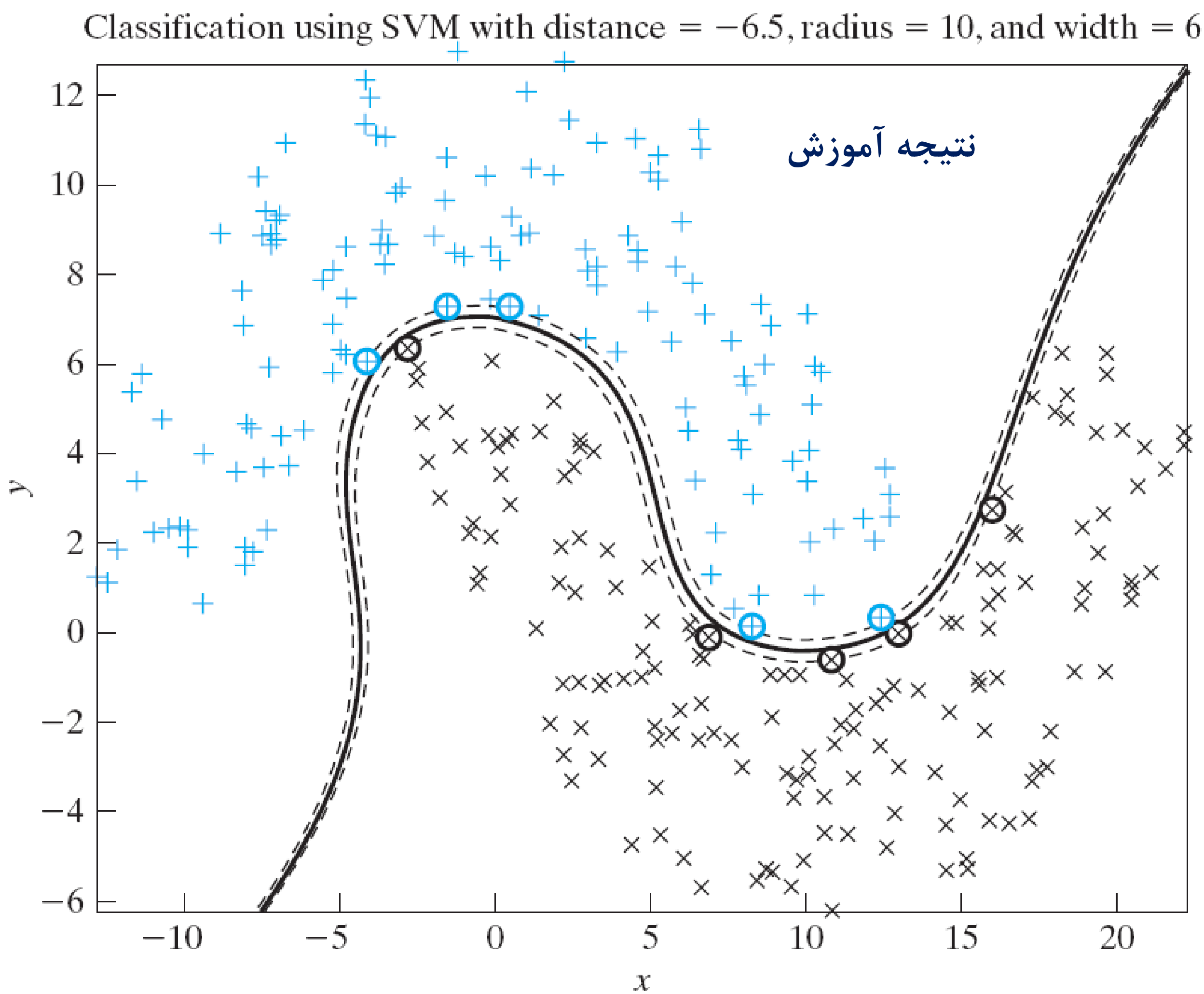


۳۰۰ داده برای آموزش
۲۰۰۰ داده برای آزمایش

$C = \infty$

ماشین بردار پشتیبان (SVM)

مثال: کلاسه‌بندی کلاس‌های ماه شکل

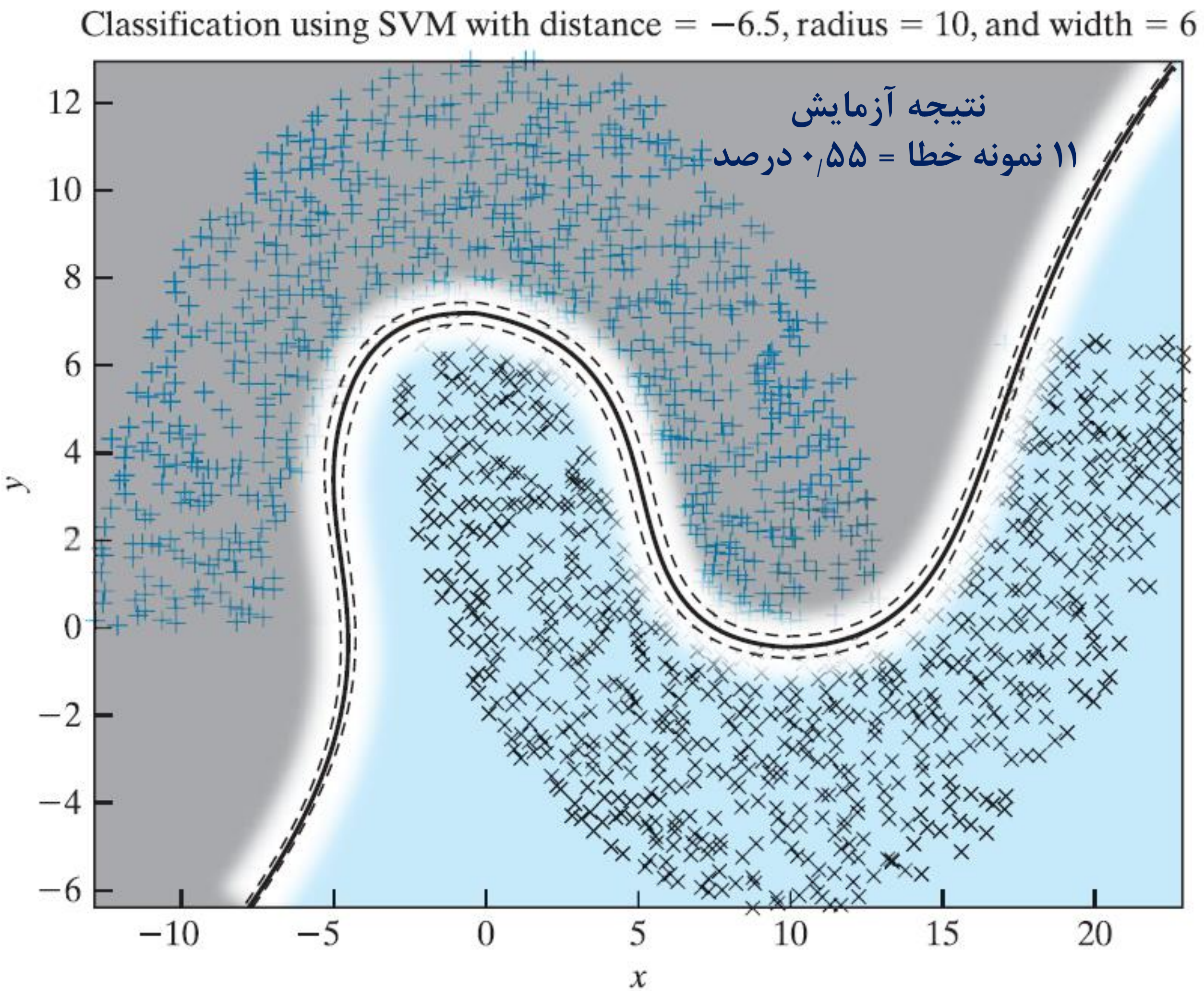


۳۰۰ داده برای آموزش
۲۰۰۰ داده برای آزمایش

$C = \infty$

ماشین بردار پشتیبان (SVM)

مثال: کلاسه‌بندی کلاس‌های ماه شکل



۳۰۰ داده برای آموزش
۲۰۰۰ داده برای آزمایش

$C = \infty$