

بسمه تعالی

بازشناسی آماری الگو



دانشگاه صنعتی خواجه نصیرالدین طوسی

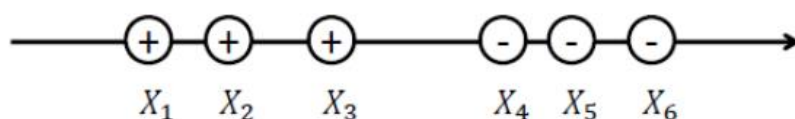
آزمایشگاه بینایی ماشین و پردازش تصاویر

تاریخ تحویل: ۱۴۰۴/۹/۲۲

تمرین سری هشتم

پژشکی

۱. (اختیاری) می‌خواهیم SVM با حاشیه تمایز نرم را روی داده‌ها تک بعدی زیر اعمال کنیم. بردارهای پشتیبان را به ازای  $c = 0$  و  $c = \infty$  تعیین نمایید.



۲. (اختیاری) داده‌های زیر را در نظر بگیرید:

$$x(1) = [1 \ -1 \ -1]^t, d(1) = -1$$

$$x(2) = [-3 \ 1 \ 1]^t, d(2) = 1$$

$$x(3) = [-3 \ 1 \ -1]^t, d(3) = -1$$

$$x(4) = [1 \ 2 \ 1]^t, d(4) = -1$$

$$x(5) = [-1 \ -1 \ 2]^t, d(5) = 1$$

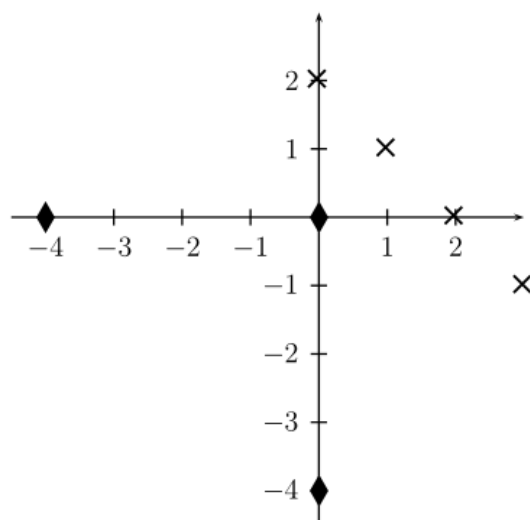
این داده‌ها خطی جداپذیر هستند. معادله صفحه جداکننده‌ای که کلاسیفایر SVM پیشنهاد می‌کند را محاسبه کنید.

۳. (اختیاری) داده‌های زیر را در نظر بگیرید:

یک SVM خطی را به شکل

$$f(\mathbf{x}) = \text{sign}(w_1x_1 + w_2x_2 + w_3)$$

با شرط  $\|\mathbf{w}\| = 1$  آموزش دهید. بردار  $\mathbf{w}$  را بیابید.



۴. (اختیاری) مسأله ۳,۲۰ کتاب Theodoridis: نشان دهید در مسأله SVM با حاشیه تمایز نرم چنانچه تابع هزینه به فرم زیر انتخاب شود:

$$J(\mathbf{w}) = \frac{1}{2} \|\mathbf{w}\|^2 + \frac{C}{2} \sum_{i=1}^N \xi_i^2$$

در این صورت مسأله می تواند مجدداً به حالت کلاس های تفکیک پذیر خطی تبدیل گردد.

۵. دو داده زیر را از دو کلاس +1 و -1 در نظر بگیرید:

$$\begin{aligned} x_1 &= (0,0), & y_1 &= +1 \\ x_2 &= (4,4), & y_2 &= -1 \end{aligned}$$

// الف) می خواهیم یک کلاسه بند SVM خطی برای این داده ها طراحی کنیم. ابتدا بردارهای پشتیبان را مشخص کرده و سپس با توجه به آنها کلاسه بند زیر را کامل کنید.

$$h(x) = \begin{cases} +1 & \text{if } \dots\dots x_1 + \dots\dots x_2 + \dots\dots \geq 0 \\ -1 & \text{otherwise} \end{cases}$$

ب) فرض کنید داده زیر نیز به داده های قبلی اضافه شود:

$$x_3 = (-1, -1), \quad y_3 = +1$$

همانند حالت قبل بردارهای پشتیبان را مشخص کرده و مرز تصمیم گیری را بیابید. آیا مرز تصمیم گیری برای این حالت متفاوت از حالت (الف) خواهد بود؟ وزن بردارهای پشتیبان در این حالت نسبت به حالت قبلی کوچکترند یا بزرگ تر؟

پ) فرض کنید داده زیر نیز به مجموعه داده های قبلی اضافه شود:

$$x_4 = (2, 2), \quad y_4 = +1$$

همانند حالت های قبل بردارهای پشتیبان را مشخص کرده و مرز تصمیم گیری را بیابید. آیا مرز تصمیم گیری برای این حالت متفاوت از حالت (ب) خواهد بود؟ وزن بردارهای پشتیبان در این حالت نسبت به حالت قبلی کوچکترند یا بزرگ تر؟

ت) فرض کنید داده زیر نیز به مجموعه داده های قبلی اضافه شود:

$$x_5 = (-3, -3), \quad y_5 = -1$$

کدام یک از کرنل های زیر قادر به جداسازی این داده ها هستند؟

- کرنل خطی:  $K(u, v) = u \cdot v$
- کرنل چند جمله ای با درجه بزرگتر مساوی ۲:  $K(u, v) = (1 + u \cdot v)^n$
- کرنل گاوسی:  $K(u, v) = \exp\left(-\frac{\|u-v\|^2}{2\sigma^2}\right)$

۶. (اختیاری) سوال امتحانی پایان ترم ۹۳

در فضای دو بعدی نقاط زیر و برچسب آنها در اختیار است:

$$\mathbf{x}_1 = \{(1,1), 1\}, \mathbf{x}_2 = \{(2,1), 1\}, \mathbf{x}_3 = \{(2,0), 1\}, \mathbf{x}_4 = \{(1,2), -1\}, \mathbf{x}_5 = \{(2,2), -1\}, \mathbf{x}_6 = \{(1,-3), -1\}$$

چنانچه از تبدیل غیرخطی زیر برای رسیدن به یک ماشین بردار پشتیبان استفاده شود

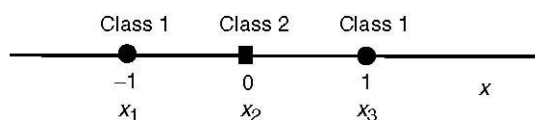
$$\boldsymbol{\varphi}(\mathbf{x}) = (x_1^2 + x_2^2, x_1 - x_2)$$

الف- بردارهای  $\boldsymbol{\varphi}_i(\mathbf{x})$   $i=1, \dots, 6$  را در فضای مخفی رسم کنید و بدون محاسبه و صرفاً بصورت شهودی بردارهای پشتیبان را بیابید.

ب- با استفاده از حل مساله دوگان مقادیر بهینه وزن و بایاس را بیابید.

ج- یک تبدیل غیرخطی معرفی کنید که داده ها را از فضای دوبعدی به فضای یک بعدی انتقال دهد طوری که دسته بندی خطی آنها امکان پذیر گردد.

۷. اختیاری (سوال امتحانی پایان ترم ۱۳۹۶) در فضای یک‌بعدی فرض کنید سه نمونه از ۲ کلاس (دایره و مربع) مطابق شکل زیر در اختیار است:

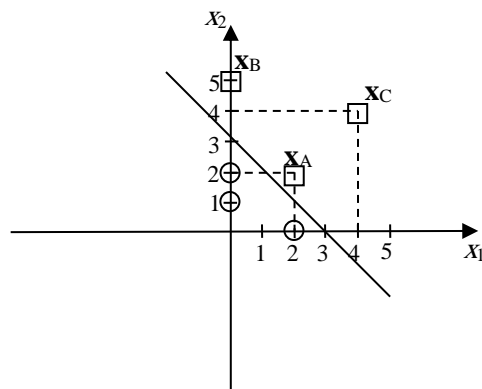


الف- نشان دهید نمونه‌های فوق بصورت خطی تفکیک‌پذیر نیستند.

ب- با استفاده از کرنل چندجمله‌ای درجه دوم  $(K(\mathbf{x}, \mathbf{x}') = (\mathbf{x}^t \mathbf{x}' + 1)^2)$  و حل مساله دوگان، معادله کلاسیفایر بهینه (مرز بین دو کلاس) را در فضای ویژگی معرفی شده بوسیله کرنل بدست آورید و آن را رسم کنید.

ج- معادله مرز بین دو کلاس در فضای اولیه (فضای یک‌بعدی) را نوشته و نشان دهید در این فضا حاشیه تمایز کلاس ۲ از حاشیه تمایز کلاس ۱ بیشتر است.

۸. اختیاری (سوال امتحانی پایان ترم ۱۳۹۷) در مساله دو کلاسه شکل زیر که در آن بردارهای ویژگی هر کلاس با علامت مشخص شده‌اند:



الف- برای مرز تصمیم‌گیری رسم شده در شکل تابع تصمیم‌گیری خطی

$g(\mathbf{x})=0$  را بنویسید و بردار عمود بر مرز  $(\mathbf{w})$  و مقدار بایاس  $(w_3)$  را بدست آورید.

ب- تابع تصمیم‌گیری بدست آمده در بند الف را به نحوی نرمالیزه کنید که داشته باشیم:

$$g'(\mathbf{x}_A) = 1$$

بردار عمود بر مرز  $(\mathbf{w}')$  و مقدار بایاس  $(w'_3)$  را بدست آورید.

ج- فرض ب را در حالت زیر انجام دهید:

$$g''(\mathbf{x}_B) = 1$$

و بردار عمود بر مرز  $(\mathbf{w}'')$  و مقدار بایاس  $(w''_3)$  را بدست آورید

د- در هریک از حالات بند ب و ج فاصله  $\mathbf{x}_B$  و  $\mathbf{x}_A$  را از مرز با استفاده از  $g'$  و  $g''$  محاسبه کنید.

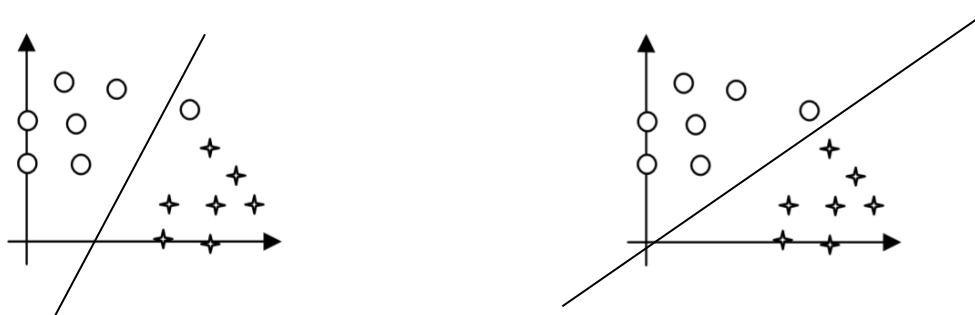
ه- با  $g''(\mathbf{x}_B) = 1$  با یک مساله بردار پشتیبان با حاشیه نرم روبرو هستیم. بردارهای پشتیبان را مشخص کنید و نشان دهید می‌توان

بردار عمود بر مرز تصمیم‌گیری بهینه را بر حسب بردارهای پشتیبان نوشت.

۹- (سوال امتحانی پایان ترم ۹۸) الف- در مساله دو کلاسه شکل زیر بردارهای ویژگی هر کلاس با علامت دایره و ستاره مشخص شده‌اند.

می‌خواهیم از SVM خطی با حاشیه نرم برای دسته‌بندی استفاده کنیم. بدون انجام حل تحلیلی و صرفاً با ارائه توجیهات لازم مطلوب است رسم

مرز تصمیم‌گیری به ازای  $C=0.05$  در شکل سمت چپ و به ازای  $C=10000$  در شکل سمت راست. در هر مورد بردارهای پشتیبان را مشخص کنید.



ب- در مساله دو کلاسه شکل زیر بردارهای ویژگی هر کلاس با علامت دایره و مربع مشخص شده‌اند. همچنین مرز تصمیم‌گیری بر اساس دسته‌بند SVM با کرنل‌ها و ضرایب جریمه ( $C$ ) مختلف نشان داده شده است. نمونه‌هایی که بصورت پررنگ نمایش داده شده‌اند، بردارهای پشتیبان هستند. مطلوب است مشخص نمودن شماره شکل مرتبط با هر یک از حالت‌های زیر:

الف- دسته‌بند SVM خطی با حاشیه تمایز نرم و  $C=0.1$

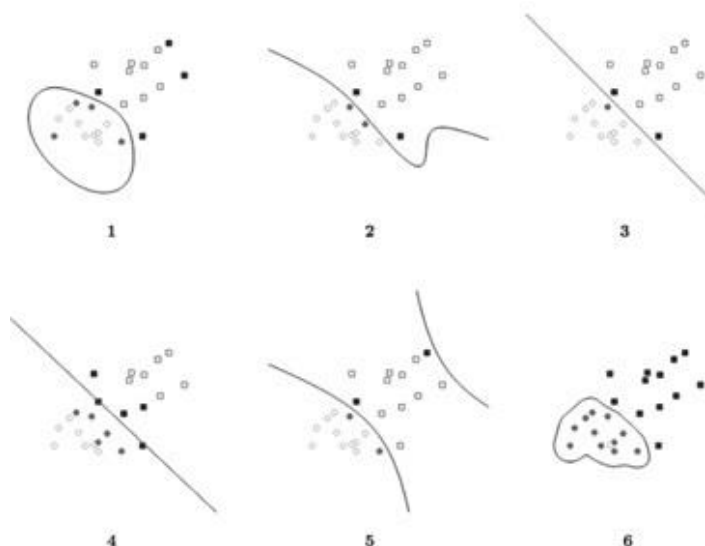
ب- دسته‌بند SVM خطی با حاشیه تمایز نرم و  $C=10$

ج- دسته‌بند SVM غیرخطی با حاشیه تمایز سخت و کرنل  $K(u, v) = u^t v + (u^t v)^2$

د- دسته‌بند SVM غیرخطی با حاشیه تمایز سخت و کرنل  $K(u, v) = \exp\left(-\frac{1}{4}\|u - v\|^2\right)$

ه- دسته‌بند SVM غیرخطی با حاشیه تمایز سخت و کرنل  $K(u, v) = \exp(-4\|u - v\|^2)$

و- هیچکدام

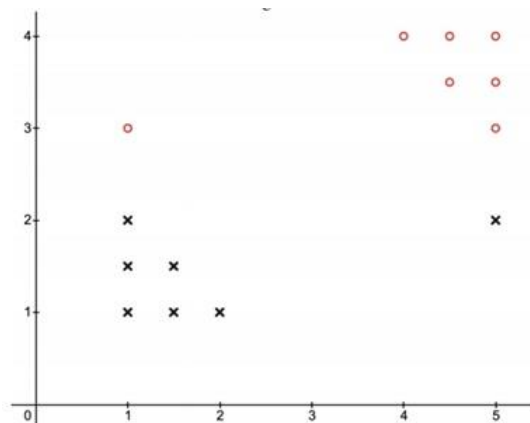


۱۰. اختیاری (سوال امتحانی پایان ترم ۱۳۹۹)

الف- در شکل زیر مطلوب است رسم مرز تصمیم‌گیری خطی SVM با حاشیه تمایز سخت

ب- با فرض مقدار محدود برای پارامتر  $C$ ، مرز تصمیم‌گیری خطی SVM با حاشیه تمایز نرم را نیز بدست آورید و رسم کنید.

ج- در دو فرض الف و ب با رسم خطوط غیر مممتد (خط چین)، حاشیه تمایز را نشان دهید.



۱۱. (اختیاری) (سوال امتحانی پایان ترم ۱۴۰۰)

یک مسئله دسته‌بندی دو کلاسه یک بعدی با نقاط  $x_1 = -2$ ,  $x_2 = 2$  و  $x_3 = 3$  با برچسب ۱- و نقطه  $x_4 = 1$  با برچسب ۱+ را در نظر بگیرید.

الف) آیا این مسئله خطی جداپذیر است؟

ب) تبدیل  $\varphi(x) = \begin{bmatrix} x \\ x^2 \end{bmatrix}$  را در نظر بگیرید. داده‌ها تحت این تبدیل به چه شکلی در می‌آیند؟ آیا در این حالت خطی جداپذیر هستند؟  
ج) مسئله را در حالت «ب» با استفاده از SVM حل کنید.

class	$x_1$	$x_2$
+	1	1
+	2	2
+	2	0
-	0	0
-	1	0
-	0	1

۹- (سوال امتحانی پایان ترم ۱۴۰۲) ماشین‌های بردار پشتیبان (SVM) مرز تصمیم‌گیری را با هدف بیشینه کردن حاشیه تمایز بدست می‌آورند. در شکل مقابل می‌خواهیم یک دسته بند SVM خطی را برای یک داده شامل ۶ نقطه بدست آوریم. داده‌های یک دسته با برچسب + و داده‌های دسته دوم با برچسب - مشخص شده‌اند.

الف - داده‌ها را در فضای دوبعدی نمایش دهید.

ب- آیا دسته‌های + و - بصورت خطی تفکیک پذیرند؟

ج- بردار وزن خط جداکننده با حاشیه تمایز بیشینه را بصورت نظری و بدون محاسبه مشخص کنید.

د- بردارهای پشتیبان و مقدار حاشیه تمایز پاسخ خود به بند قبل را مشخص کنید.

ه- چنانچه یکی از بردارهای پشتیبان حذف شود، آیا حاشیه تمایز افزایش می‌یابد؟ چرا؟

و- حذف کدامیک از بردارهای پشتیبان موجب افزایش حاشیه تمایز خواهد شد؟

ز- چنانچه یک نمونه برچسب‌دار به نمونه‌های جدول فوق اضافه شود، آیا حاشیه تمایز کاهش می‌یابد؟ چرا؟

ح- اضافه شدن یک نمونه برچسب‌دار چه تغییری (افزایش/بدون تغییر/کاهش) در تعداد بردارهای پشتیبان بوجود می‌آید؟ با ارائه مثال در هر مورد پاسخ خود را توضیح دهید.

۱۰- (پایان ترم ۱۴۰۳) فرض کنید که داده‌های آموزش زیر در فضای دوبعدی در اختیار هستند و می‌خواهیم با استفاده از kernel-SVM آنها را تفکیک کنیم.

برچسب	$x_1$	$x_2$
+	0.2	0.4
+	0.4	0.8
+	0.8	0.4
+	2	4.0
-	0.4	0.4
-	0.8	0.8

اگر داده‌های  $x$  را به فضای  $\phi(x)$  نگاشت دهیم و تابع کرنل به صورت زیر باشد:

$$k(x, x') = \phi(x)^T \phi(x') = \frac{x^T x'}{\|x\| \|x'\|}$$

الف) داده‌ها را رسم کنید و مشخص کنید که آیا تفکیک‌پذیر خطی هستند یا خیر.

ب) بردار ویژگی  $\phi(x)$  مربوط به این کرنل را به دست آورید و نگاشت داده‌های آموزش را در این فضا رسم کنید. آیا داده‌ها در فضای جدید تفکیک‌پذیر خطی هستند؟

ج) بردارهای پشتیبان را در فضای جدید مشخص کنید. خط جداکننده در فضای جدید را به دست آورید. (در صورتی که، بدون حل معادلات، می‌توانید خط را پیدا کنید، روش خود را توضیح دهید).

د) با استفاده از نتایج به دست آمده در استفاده از کرنل، مرز تصمیم‌گیری در فضای اولیه‌ی داده‌های آموزش را به صورت تقریبی رسم نمایید (با توجه به تفکیک‌پذیر بودن یا نبودن داده‌ها در فضای ویژگی جدید و فرم کلی معادله).

موفق باشید