۱- در روش رگرسیون خطی تابع هزینه را به صورت $J(\underline{w}) = E\left\{(y - \underline{w}^T \underline{x})^2\right\}$ تعریف می‌کنیم. اولا این تابع هزینه را ساده کنید. ثانیا اگر داده‌ها نویزی به صورت $\underline{x} + \underline{\varepsilon}$ باشد (که بردار نویز با متوسط صفر و ماتریس کواریانس $\sigma^2 I$ و مستقل از داده فرض می‌شود)، تابع هزینه را برحسب تابع هزینه حالت بدون نویز بدست آورید.

۲- داده‌های زیر به همراه برچسب‌هایشان را داریم:

$$\left\{\left(\begin{bmatrix}1\\2\end{bmatrix}, +1\right), \left(\begin{bmatrix}2\\1\end{bmatrix}, +1\right), \left(\begin{bmatrix}-1\\-2\end{bmatrix}, -1\right), \left(\begin{bmatrix}-2\\-1\end{bmatrix}, -1\right)\right\}$$

الف) با استفاده از الگوریتم پرسپترون و اندازه گام ۰٫۱ مرز تصمیم‌گیری را بدست آورده و آن را رسم کنید. ( $\underline{w}_0 = \begin{pmatrix}0\\0\\1\end{pmatrix}$ )

ب) با استفاده از الگوریتم پاداش و جریمه و اندازه گام ۰٫۱ مرز تصمیم‌گیری را بدست آورده و آن را رسم کنید. ( $\underline{w}_0 = \begin{pmatrix}0\\0\\1\end{pmatrix}$ )

پ) با روش Hard SVM معادله مرز تصمیم‌گیری و دو حاشیه را بدست آورده و آن را رسم کنید. معادلات و شرایط KKT را به طور کامل بنویسید و یک دسته جواب برای ضرایب لاگرانژ تعیین کنید. آیا این دسته ضرایب یکتا هستند؟ اگر خیر، یک دسته دیگر هم تعیین کنید.

۳- الف) برای چهار داده $AND$، معادله خط تصمیم‌گیری را با استفاده از الگوریتم $LS$ $(SES)$ بدست آورید.

ب) برای چهار داده $OR$، معادله خط تصمیم‌گیری را با استفاده از الگوریتم $LS$ $(SES)$ بدست آورید.

۴- ضرایب بردار وزن در روش Logistic Discrimination برای حالت دو کلاس هم‌احتمال با توزیع گوسی ویژگی‌ها و ماتریس کواریانس یکسان بدست آورید. جواب را با طبقه‌بندی کننده بیز مقایسه کنید.

۵- مسئله ۱۴ از فصل سوم را با فرض‌های زیر حل کنید. بعد ویژگی ۱ بوده و از کلاس اول $N_1$ و از کلاس دوم $N_2$ داده داریم و بجای روش آماری $MSE$ از روش جبری $SES$ (صفحه ۲۷ فایل مبحث سوم) مسئله را حل می‌کنیم. برچسب داده‌های کلاس اول را $-\dfrac{N_1 + N_2}{N_1}$ و کلاس دوم را $-\dfrac{N_1 + N_2}{N_2}$ می‌گیریم. در روش $LDA$، .میانگین کلاس‌ها و واریانس را از روی داده‌های آموزش تخمین می‌زنیم.

۶- الف) در الگوریتم $SVM$ برای حالت خطی جدایی‌پذیر، آیا تمام بردارهایی که روی دو حاشیه هستند، در ساختن ابرصفحه مرز تصمیم‌گیری (یعنی بردار $\underline{W}$) شرکت می‌کنند؟ چرا؟

ب) در الگوریتم $C-SVM$، نشان دهید بردارهایی که در ساختن ابر صفحه مرز تصمیم‌گیری (یعنی بردار $\underline{W}$) شرکت می‌کنند در یکی از چهار گروه هستند: ۱) روی دو حاشیه هستند و درست طبقه‌بندی می‌شوند، ۲) در باند بین دو حاشیه هستند و درست طبقه‌بندی می‌شوند، ۳) روی مرز تصمیم‌گیری هستند، ۴) داده‌هایی که غلط طبقه‌بندی می‌شوند (ممکن است روی دو حاشیه هم باشند).

سپس نشان دهید از بین چهار دسته فوق، دسته اول کمترین تاثیر را در ساختن بردار $\underline{W}$ دارند و سه گروه دیگر وزن یکسانی دارند.

۷- می‌خواهیم با در نظر گرفتن داده‌های زیر و با کمک رگرسیون خطی، یک تخمین غیرخطی با مدل زیر بدست آوریم. ضرایب مدل را بدست آورید.

$$Y = f(x) = e^{a+bx}$$

| | | | | | | |
|---|---|---|---|---|---|---|
| x | 0.00 | 1.00 | 2.50 | 3.51 | 4.20 | 7.00 |
| y | 1.00 | 1.52 | 2.72 | 5.10 | 6.60 | 7.00 |

۸- در مسئله $SVM$ برای حالت خطی جدایی ناپذیر فرض کنید: $I(\eta_i) = \frac{1}{2}(\eta_i)^2$ (که به آن $L2-SVM$ می‌گویند)

الف) تابع لاگرانژین را بنویسید.

ب) سپس شرایط $KKT$ را بنویسید.

پ) فرم دوگان $Wolfe$ را نوشته و سپس فرم ساده شده آن را بنویسید.

ت) نشان دهید حذف قید $\eta_i \geq 0$ در جواب تغییری ایجاد نمی‌کند.

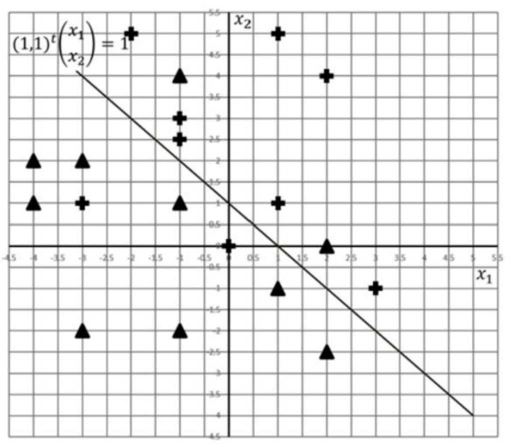
۹- برای روش $ECOC$ با مقادیر $M = 4, L = 6$ و برای ماتریس داده شـــده در صـــفحه ۱۲۸ کتاب، بردار $(-1,-1,-1,-1,-1,-1)$ برای یک داده تست بدست آمده است. برچسب این بردار را تعیین کنید.

۱۰- فرض کنید برای داده‌های شکل زیر و برای مسئله $C-SVM$ معادله مرز بدست آمده است و در شکل رسم شده است.

الف) معادله دو حاشیه را بنویسید.

ب) فرض کنید $\eta_{(x_1,x_2)}$ متغیر slack برای $\boldsymbol{x} = (x_1, x_2)$ باشد. در آن صورت برای هر یک از گزاره‌های زیر کدام یک از علائم $=$، $<$، $>$ را مناسب می‌دانید؟

- $\eta_{(2,4)} \square \eta_{(2,0)}$
- $\eta_{(-1,1)} \square \eta_{(-1,-2)}$
- $\eta_{(-3,1)} \square \eta_{(-1,2.5)}$
- $\eta_{(2,4)} \square \eta_{(-1,-2)}$




پ) برای هر داده آموزشی یکی شرایط زیر بوجود می‌آید:

۱. نمونه مورد نظر قطعا یک بردار پشتیبان است.

۲. نمونه مورد نظر قطعا یک بردار پشتیبان نیست.

۳. نمی‌توان با توجه به شرایط موجود گفت که یک نمونه می‌تواند یک بردار پشتیبان باشد یا خیر.

برای هر کدام از نمونه‌های زیر تعیین کنید کدام شرط از شرایط ۱ تا ۳ بالا برای آن محتمل‌تر است.

- $\boldsymbol{x} = (2, 4)$
- $\boldsymbol{x} = (1, 1)$
- $\boldsymbol{x} = (2, 0)$

۱۱- فرض کنید تابع هدف $J(\underline{w}) = \underline{w}^T A \underline{w}$ را داریم که $A$ یک ماتریس متقارن بوده که با استفاده از بردارها و مقادیر ویژه آن داریم:
$A = U \Lambda U^T$

الف) با استفاده از گرادیان نزولی رابطه بازگشتی برای محاسبه بردار $\underline{w}$ را بنویسید.
ب) شرایط همگرایی الگوریتم را با استفاده از قسمت قبل بدست آورید.

۱۲- پایگاه داده زیر را در نظر بگیرید:

$$\left\{\left(\begin{bmatrix}2\\3\end{bmatrix},1\right),\left(\begin{bmatrix}3\\2\end{bmatrix},1\right),\left(\begin{bmatrix}3\\3\end{bmatrix},-1\right),\left(\begin{bmatrix}1\\1\end{bmatrix},-1\right),\left(\begin{bmatrix}-1\\1\end{bmatrix},-1\right)\right\}$$

فرض کنید برای پایگاه داده فوق یک $C-SVM$ با C=10 آموزش داده‌اید.
الف) دو حاشیه و مرز را حدس زده و روی شکل رسم کنید.
ب) کدامیک از بردارهای زیر یک بردار $\lambda$ معتبر برای مساله فوق است (توجه کنید که تنها یکی از بردارها معتبر است و برای هر کدام از بردارهای غیرمعتبر باید یک دلیل بیاورید).

$$\lambda^{(1)} = \begin{bmatrix}6.89\\6.89\\10\\3.78\\0\end{bmatrix}, \lambda^{(2)} = \begin{bmatrix}3.78\\3.78\\5\\3.78\\0\end{bmatrix}, \lambda^{(3)} = \begin{bmatrix}6.89\\6.89\\10\\0\\3.78\end{bmatrix}, \lambda^{(4)} = \begin{bmatrix}0\\3.78\\10\\6.89\\0\end{bmatrix}, \lambda^{(5)} = \begin{bmatrix}6.89\\6.89\\10\\6.89\\0\end{bmatrix}$$

پ) بر اساس بردار معتبری که در قسمت الف به دست آورده‌اید، بردار وزن و بایاس مرز تصمیم‌گیری را تعیین کنید.

۱۲- در یک مسئله رگرسیون خطی داریم:

$y = \underline{w}^T \underline{x},\ \underline{x} \in \Re^L\ \ y \in \Re$

$$X = [\underline{x}_1\ \ \underline{x}_2\ \ \cdots\ \ \underline{x}_N],\quad \underline{y} = \begin{pmatrix}y_1\\y_2\\\vdots\\y_N\end{pmatrix}\quad \underline{w} = \begin{pmatrix}w_1\\w_2\\\vdots\\w_L\end{pmatrix}$$

الف) اگر رگرسیون را فقط بر روی ویژگی jام انجام دهیم، نشان دهید $w_j = \dfrac{X_j \underline{y}}{X_j X_j^T}$ که $X_j$ سطر jام ماتریس داده‌ها است.
ب) فرض کنید ویژگی‌ها مستقل هستند (یعنی سطرهای ماتریس داده‌ها مستقل هستند). ثابت کنید که پارامترهای بهینه از آموزش رگرسیون بر روی همه ویژگی‌ها (که در درس بدست آمده است) با پارامترهای بهینه حاصل از آموزش روی هر ویژگی به طور مستقل یکسان است.
پ) فرض کنید $y = \underline{w}^T \underline{x} + w_0$. اگر رگرسیون را فقط بر روی ویژگی jام انجام دهیم، $w_j, w_0$ را بدست آورید.