Subject:
Year. Month. Date. ( )

۴۰۰۱۰۴۸۹۱ عرفان جعفری

۱_ الف) : $K_1(x,z)+K_2(x,z)$

$K_1 = \phi_1^T(x)\,\phi_1(z)$
$K_2 = \phi_2^T(x)\,\phi_2(z)$
$\Rightarrow K_1(x,z)+K_2(x,z) = \phi_1^T(x)\,\phi_1(z) + \phi_2^T(x)\,\phi_2(z) = \left\langle \underbrace{\begin{bmatrix}\phi_1(x)\\ \phi_2(x)\end{bmatrix}}_{\phi'(x)}, \underbrace{\begin{bmatrix}\phi_1(z)\\ \phi_2(z)\end{bmatrix}}_{\phi'(z)} \right\rangle$

$\Longrightarrow K(x,z) = \phi'^T(x)\,\phi'(z) \rightarrow$ کرنل معتبر است

ب) : $aK_1(x,z)$

$aK_1(x,z) = a\,\phi^T(x)\,\phi(z) = \phi^T(\sqrt{a}\,x)\,\phi(\sqrt{a}\,z) = \langle \underbrace{\sqrt{a}\,\phi(x)}_{\phi'(x)}, \underbrace{\sqrt{a}\,\phi(z)}_{\phi'(z)} \rangle$
$= \phi'^T(x)\,\phi'(z) \longrightarrow$ کرنل معتبر است

ج) : $K_2(f(x), f(z))$

$\phi^T(f(x))\,\phi(f(z)) = \langle \phi(f(x)), \phi(f(z)) \rangle$ $\phi(f(x)) = \phi'(x)$ و $\phi(f(z)) = \phi'(z)$

$= \langle \phi'(x), \phi'(z) \rangle = \phi'^T(x)\,\phi(z) \longrightarrow$ کرنل معتبر است

۲_ الف) هنگامی که Regularization Term مقدار بزرگی می‌گیرد بخش رگرسیون ساده وزن کمتری به خود اختصاص داده و مدل تلاش می‌کند با کمتر کردن w_i ها، جریمه‌ای که در اثر بخش تعدیلگر دریافت می‌کند را کم کند، یعنی دسته‌بندی بدون توجه به x_2 انجام می‌شود و از آن جا که نمونه‌ها به کمک فیچر x_2 قابل جداسازی هستند، خطا زیاد نمی‌شود و ثابت می‌ماند.

ب) با توجه به توضیحات بخش قبل، مقدار w_1 کوچک شده و چون داده‌ها بر حسب x_2 به تنهایی قابل دسته‌بندی نیستند، خطا افزایش می‌یابد.

ج) با حذف w_0 از عبارت رگرسیون، دقت مدل پایین می‌آید و خطا افزایش پیدا می‌کند.

Subject:
Year. Month. Date. ( )

د) چون مدل به دنبال کم کردن $\|w_1\|$ و $\|w_2\|$ است در عبارت رگرسیون، تنها پارامتر موثر w_0 برای classification خواهد بود. یعنی مدل به کمک $\frac{1}{1+e^{-w_0}}$ نمونه‌ها را دسته بندی می‌کند که چون با توجه به شکل، تعداد نمونه‌ها یکسان است، می‌توان گفت $P(y=1|x) = P(y=0|x)$ و هر دو برابر $\frac{1}{2}$ هستند؛ بنابراین:

$$\frac{1}{1+e^{-w_0}} = \frac{1}{2} \Rightarrow e^{-w_0} = 1 \Rightarrow \underline{w_0 = 0}$$

ه) با توجه به استدلال قسمت قبل، می‌توان گفت

$$P(y=1|x) > \frac{1}{2} \rightarrow \frac{1}{1+e^{-w_0}} > \frac{1}{2} \rightarrow e^{-w_0} < 1 \rightarrow \underline{w_0 > 0}$$

بنابراین w_0 می‌تواند مقادیر بزرگتر از ۰ را به خود اختصاص دهد.

۳_ الف) از بسط تیلور برای بازنویسی کرنل گوسی استفاده می‌کنیم:

$$K(x,x') = e^{-\frac{\|x-x'\|^2}{2\sigma^2}} = \sum_0^\infty \underbrace{\frac{(-1)^n}{n!\,2\sigma^2}}_{c_n} (\|x-x'\|^2)^n = \sum_0^\infty c_n (\|x-x'\|^2)^n$$

چون n محدودیتی ندارد، آن را می‌توان تا بی‌نهایت افزایش داد که بردار ویژگی ابعادش نامتناهی باشد، همچنین چون کرنل است آن را می‌توان بصورت ضرب داخلی این بردار ویژگی‌ها با بعد نامتناهی نوشت.

ب) چون A متقارن و مثبت نیمه معین است، آن را بصورت زیر تجزیه می‌کنیم:

$$A = U^T \Lambda U = U^T \Lambda^{\frac{1}{2}} \Lambda^{\frac{1}{2}} U = (\Lambda^{\frac{1}{2}} U)^T \Lambda^{\frac{1}{2}} U \quad \text{(eigenvalue decomposition)}$$

ماتریس قطری مقادیر ویژه که همگی نامنفی هستند

Subject:

Year. Month. Date. ( )

چون ماتریس Λ قطری بوده و مقادیر روی قطر (مقادیر ویژه A) همگی نامنفی هستند، آن را به صورت $(\Lambda^{\frac{1}{2}})^T\Lambda^{\frac{1}{2}} = \Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}$

بازنویسی کردیم؛ حال اگر در معادله x^TAy جایگذاری کنیم داریم:

$$K(x,y) = x^TAy = x^T(\Lambda^{\frac{1}{2}}U)^T\Lambda^{\frac{1}{2}}Uy = (\Lambda^{\frac{1}{2}}Ux)^T\Lambda^{\frac{1}{2}}Uy \quad , \quad \phi(x) = \Lambda^{\frac{1}{2}}Ux$$

$$\phi(y) = \Lambda^{\frac{1}{2}}Uy$$

$$\Longrightarrow K(x,y) = \phi(x)^T\phi(y)$$

بنابراین یک کرنل معتبر است.