Support Vector Machines
Homework 04

Due March 08, 2016

Robert Brown

1 Problem 7.1

Given $\phi(x) = x^2 - x$ and $g(x) = x - 3$ we construct $L(\alpha, w, b)$ according to $L(\alpha, w, b) = \phi(x) - \sum_{i=1}^{\ell} \alpha_i g_i(x)$ and find $L(\alpha, w, b) = \phi(x) - \alpha g(x) = (x^2 - x) - \alpha(x - 3)$.

We know that L is minimized with respect to α when $\frac{\partial L}{\partial \alpha} = 0$ and similarly maximized with respect to x when $\frac{\partial L}{\partial x} = 0$.

$$
\begin{aligned}
\frac{\partial L}{\partial x} &= (2x - 1) - \alpha = 0 \\
\frac{\partial L}{\partial \alpha} &= 3 - x = 0 \\
x &= 3 \\
\alpha &= 5
\end{aligned}
$$

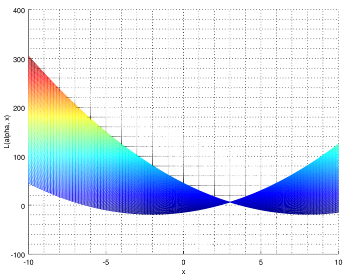

Figure 1: global minima (x)

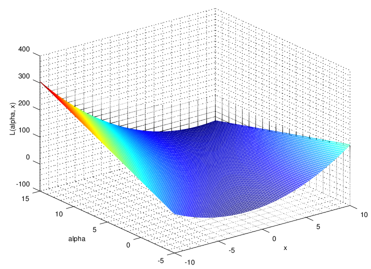

Figure 2: Solution space

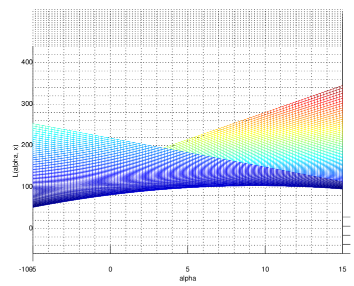

Figure 3: global maxima (α)

2 Problem 7.3

$$L(\boldsymbol{\alpha}, \mathbf{w}, b) = \tfrac{1}{2}\mathbf{w} \cdot \mathbf{w} - \sum_{i=1}^{\ell} \alpha_i y_i \mathbf{w} \cdot \mathbf{x_i} + b \sum_{i=1}^{\ell} \alpha_i y_i + \sum_{i=1}^{\ell} \alpha_i$$

Equation 7.41

$$\frac{\partial L(\boldsymbol{\alpha}, \mathbf{w}, b)}{\partial \mathbf{w}} = \frac{\partial}{\partial \mathbf{w}}(\tfrac{1}{2}\mathbf{w} \cdot \mathbf{w}) - \frac{\partial}{\partial \mathbf{w}}(\sum_{i=1}^{\ell} \alpha_i y_i \mathbf{w} \cdot \mathbf{x_i}) + \cancel{\frac{\partial}{\partial \mathbf{w}}(b \sum_{i=1}^{\ell} \alpha_i y_i)} + \cancel{\frac{\partial}{\partial \mathbf{w}}(\sum_{i=1}^{\ell} \alpha_i)} = \frac{1}{2}\frac{\partial}{\partial \mathbf{w}}(\mathbf{w} \cdot \mathbf{w}) - \sum_{i=1}^{\ell} \alpha_i y_i \frac{\partial}{\partial \mathbf{w}}(\mathbf{w} \cdot \mathbf{x_i})$$

$$= \frac{1}{2}\begin{bmatrix} \frac{\partial}{\partial w_1} & \frac{\partial}{\partial w_2} & \cdots & \frac{\partial}{\partial w_n} \end{bmatrix} \mathbf{w} \cdot \mathbf{w} - \begin{bmatrix} \frac{\partial}{\partial w_1} & \frac{\partial}{\partial w_2} & \cdots & \frac{\partial}{\partial w_n} \end{bmatrix} \sum_{i=1}^{\ell} \alpha_i y_i \mathbf{w} \cdot \mathbf{x_i}$$

$$= \frac{1}{2}\sum_{i=1}^{n} \frac{\partial}{\partial w_i}(w_i^2)\hat{e_i} - \sum_{i=1}^{\ell} \left(\alpha_i y_i \sum_{k=1}^{n} \frac{\partial}{\partial w_k}(w_k x_k)\hat{e_k} \right) = \sum_{i=1}^{n} w_i \hat{e_i} - \sum_{i=1}^{\ell} \left(\alpha_i y_i \sum_{k=1}^{n} x_k \hat{e_k} \right)$$

Since $\mathbf{w}$ and $\mathbf{x}$ are both in the same bector space ($\mathbf{R}^n$), they share the same standard basis vectors $(\hat{e_1}, \hat{e_2}, \ldots, \hat{e_n})$ this finally simplifies to $\mathbf{w} - \sum_{i=1}^{\ell} \alpha_i y_i \mathbf{x_i}$.

Equation 7.43

$$\frac{\partial L(\boldsymbol{\alpha}, \mathbf{w}, b)}{\partial \mathbf{w}} = \cancel{\frac{\partial}{\partial b}(\tfrac{1}{2}\mathbf{w} \cdot \mathbf{w})} - \cancel{\frac{\partial}{\partial b}\left(\sum_{i=1}^{\ell} \alpha_i y_i \mathbf{w} \cdot \mathbf{x}\right)} + \frac{\partial}{\partial b}\left(b \sum_{i=1}^{\ell} \alpha_i y_i\right) + \cancel{\frac{\partial}{\partial b}\left(\sum_{i=1}^{\ell} \alpha_i\right)} = \frac{\partial}{\partial b}\left(b \sum_{i=1}^{\ell} \alpha_i y_i\right) = \sum_{i=1}^{\ell} \alpha_i y_i$$

3 Problem 7.5

$$L(\boldsymbol{\alpha}, \mathbf{w}, b) = \tfrac{1}{2}\mathbf{w} \cdot \mathbf{w} - \sum_{i=1}^{\ell} \alpha_i y_i \mathbf{w} \cdot \mathbf{x} + b \sum_{i=1}^{\ell} \alpha_i y_i + \sum_{i=1}^{\ell} \alpha_i$$

By equation 7.43, we have $\sum_{i=1}^{\ell} \alpha_i y_i = 0$, and $\boldsymbol{w}^* = \sum_{i=1}^{\ell} \alpha_i y_i \mathbf{x_i}$ by equation 7.42.

Now, consider the quantity $\boldsymbol{w}^* \boldsymbol{w}^*$:

$$\boldsymbol{w}^* \boldsymbol{w}^* = \left(\sum_{i=1}^{\ell} \alpha_i y_i \mathbf{x_i} \right) \cdot \left(\sum_{j=1}^{\ell} \alpha_j y_j \mathbf{x_j} \right) = \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \alpha_j \alpha_i y_j y_i \mathbf{x_j} \cdot \mathbf{x_i}$$

And substitue these values into our laplacian:

$$\begin{aligned}
\tfrac{\partial L(\boldsymbol{\alpha}, \mathbf{w}, b)}{\partial \mathbf{w}} &= \tfrac{1}{2}\mathbf{w} \cdot \mathbf{w} - \sum_{i=1}^{\ell} \alpha_i y_i \mathbf{w} \cdot \mathbf{x} + b \cancel{\sum_{i=1}^{\ell} \alpha_i y_i} + \sum_{i=1}^{\ell} \alpha_i \\
&= \tfrac{1}{2} \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \alpha_j \alpha_i y_j y_i \mathbf{x_j} \cdot \mathbf{x_i} - \sum_{i=1}^{\ell} \alpha_i y_i \left(\sum_{j=1}^{\ell} \alpha_j y_j \mathbf{x_j} \right) \cdot \mathbf{x_i} + \sum_{i=1}^{\ell} \alpha_i \\
&= \tfrac{1}{2} \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \alpha_j \alpha_i y_j y_i \mathbf{x_j} \cdot \mathbf{x_i} - \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \alpha_j \alpha_i y_j y_i \mathbf{x_j} \cdot \mathbf{x_i} + \sum_{i=1}^{\ell} \alpha_i \\
&= \sum_{i=1}^{\ell} \alpha_i - \frac{1}{2} \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \alpha_j \alpha_i y_j y_i \mathbf{x_j} \cdot \mathbf{x_i}
\end{aligned}$$