# Lecture 6

Chiranjib Bhattacharyya

chibha@chalmers.se

February 5, 2012

# Agenda

- Support vector machines

## Support Vector Machines

$$min_{w,b} \frac{1}{2} \|w\|^2$$

$$y_i(w^\top x_i + b) \geq 1$$

# Some Interesting facts

$$min_x f(x) \ \ s.t. g_i(x) \geq 0 i = 1, \ldots, n \quad \text{(P)}$$

$$\text{Lagrangian} L(x, \alpha) = f(x) - \sum_{i=1}^{n} \alpha_i g_i(x)$$

$\alpha_i \geq 0$
K.K.T conditons

$$\frac{\partial L}{\partial x} = 0 \ g_i(x) \geq 0$$

$$\alpha_i g_i(x) = 0, \alpha_i \geq 0$$

# Details

$$L(w, b, \alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} \alpha_i (y_i (w^\top x_i + b) - 1)$$

## KKT conditions

$$\frac{\partial L}{\partial w} = w - \sum_{i=1}^{n} \alpha_i y_i x_i$$

$$\frac{\partial L}{\partial b} = -\sum_{i=1}^{n} \alpha_i y_i$$

$$\alpha_i (y_i (w^\top x_i + b) - 1) = 0 \quad y_i (w^\top x_i + b) \geq 1$$

# SVMs

At optimality $\alpha_i \geq 0$

$$w = \sum_{i=1}^{n} \alpha_i y_i x_i$$

$$b = y_i - w^\top x_i \quad \alpha_i > 0$$

## SVM Dual

$$\max_{\alpha} \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j x_i^\top x_j$$

$$\alpha_i \geq 0 \quad \sum_{i=1}^{n} \alpha_i y_i = 0$$

# Geometric understanding

- The class $y = 1$ lies in the halfspace

$$w^\top x + b \geq 1$$

- The class $y = -1$ lies in the halfspace

$$w^\top x + b \leq -1$$

- The support vectors either lie in $w^\top x + b = 1$ or $w^\top x + b = -1$

## SVMs(Non-Separable case)

$$min_{w,b,\xi_i} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n} \xi_i$$

$$y_i(w^\top x_i + b) \geq 1 - \xi_i \;\; \xi_i \geq 0 \; i = 1, \ldots, n$$

# Details

$$L(w,b,\xi,\alpha,\mu) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i$$

$$-\sum_{i=1}^{n}\alpha_i(y_i(w^\top x_i + b) - 1 + \xi_i) - \sum_{i=1}^{n}\mu_i\xi_i$$

## KKT conditions

$$\frac{\partial L}{\partial w} = w - \sum_{i=1}^{n}\alpha_i y_i x_i$$

$$\frac{\partial L}{\partial b} = -\sum_{i=1}^{n}\alpha_i y_i \quad \frac{\partial L}{\partial \xi_i} = C - \alpha_i - \mu_i$$

$$\alpha_i(y_i(w^\top x_i + b) - 1 + \xi_i) = 0, \;\; y_i(w^\top x_i + b) \geq 1 - \xi_i, \mu_i\xi_i = 0$$

# SVMs(Non-separable case)

At optimality $\alpha_i \geq 0$

$$w = \sum_{i=1}^{n} \alpha_i y_i x_i$$

$$b = y_i - w^\top x_i \forall C > \alpha_i > 0$$

$$\xi_i > 0 \Leftarrow \alpha_i = C$$

## SVM Dual

$$\max_{\alpha} \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j x_i^\top x_j$$

$$C \geq \alpha_i \geq 0 \quad \sum_{i=1}^{n} \alpha_i y_i = 0$$