

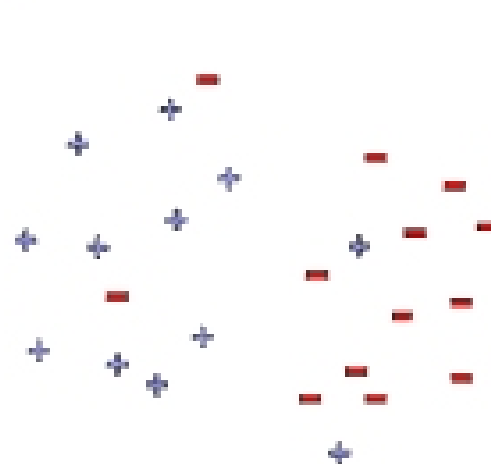
SVMs, Duality and the Kernel Trick

Machine Learning – 10701/15781
Carlos Guestrin
Carnegie Mellon University
October 21st, 2009

©Carlos Guestrin 2005-2009

1

SVMs reminder


$$\begin{aligned} \text{minimize}_{\mathbf{w}, b} \quad & \mathbf{w} \cdot \mathbf{w} + C \sum_j \xi_j \\ & (\mathbf{w} \cdot \mathbf{x}_j + b) y_j \geq 1 - \xi_j, \quad \forall j \\ & \xi_j \geq 0, \quad \forall j \end{aligned}$$

©Carlos Guestrin 2005-2009

2

Today's lecture

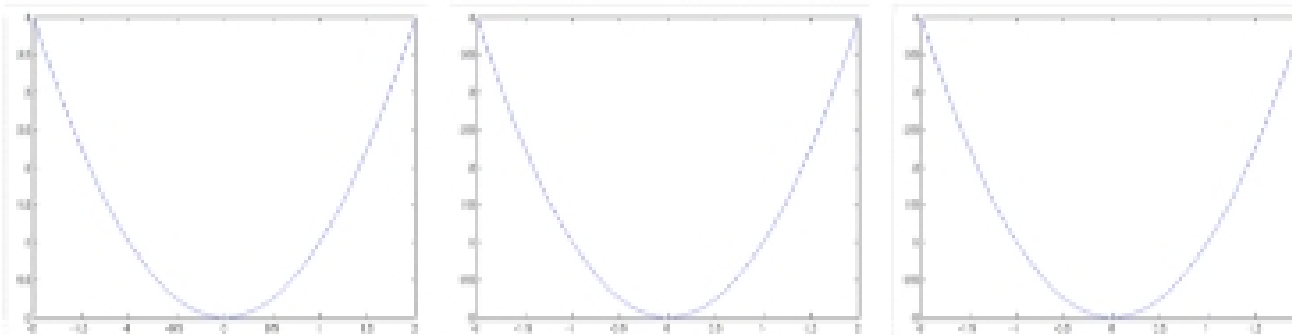
- Learn one of the most interesting and exciting recent advancements in machine learning
 - The “kernel trick”
 - High dimensional feature spaces at no extra cost!
- But first, a detour
 - Constrained optimization!

©Carlos Guestrin 2009-2020

2

Constrained optimization

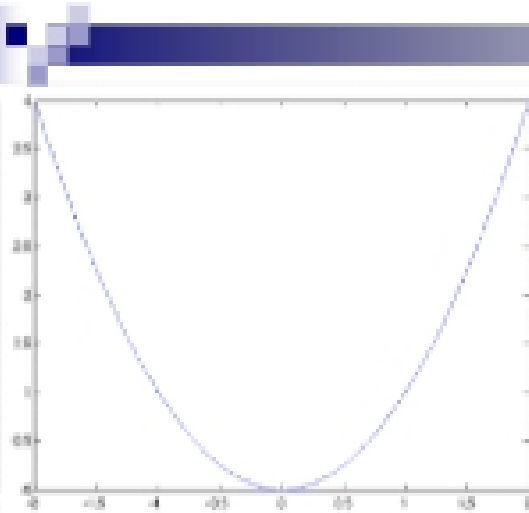
$$\begin{aligned} \min_x \quad & x^2 \\ \text{s.t.} \quad & x \geq b \end{aligned}$$



©Carlos Guestrin 2009-2020

3

Lagrange multipliers – Dual variables



$$\min_x x^2$$
$$\text{s.t. } x \geq b$$

Moving the constraint to objective function

Lagrangian:

$$L(x, \alpha) = x^2 - \alpha(x - b)$$
$$\text{s.t. } \alpha \geq 0$$

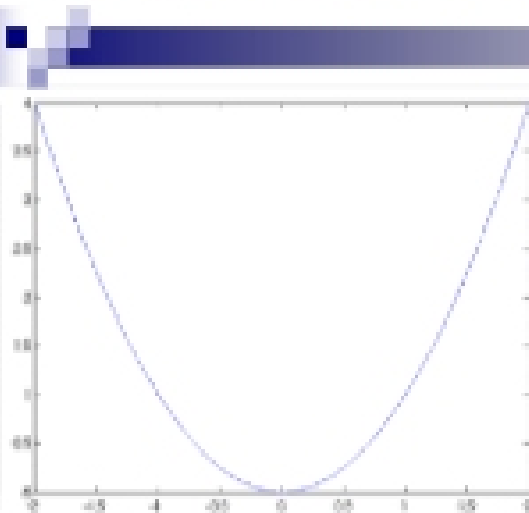
Solve:

$$\min_x \max_{\alpha} L(x, \alpha)$$
$$\text{s.t. } \alpha \geq 0$$

©Carlos Queiroz 2009-2020

1

Lagrange multipliers – Dual variables



Solving: $\min_x \max_{\alpha} x^2 - \alpha(x - b)$

$$\text{s.t. } \alpha \geq 0$$

©Carlos Queiroz 2009-2020

1