CS 383: Machine Learning

Prof Adam Poliak

Fall 2024

11/06/2024

Lecture 21

SVM

Idea: put arbitrary constraint on functional margin

$$\boxed{\hat{\gamma} = 1}$$

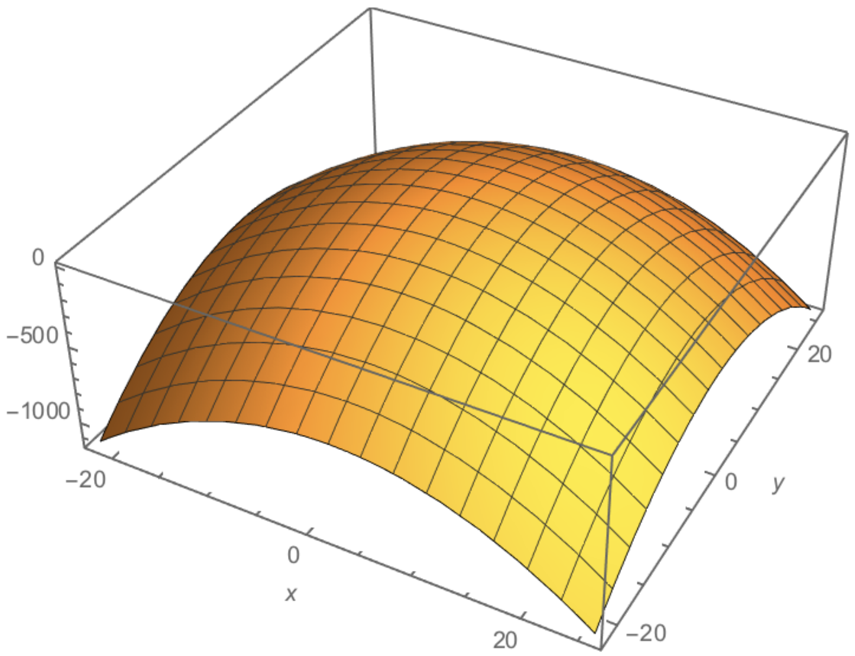
$$\min_{\vec{w},b} \quad \frac{1}{2}\|\vec{w}\|^2$$

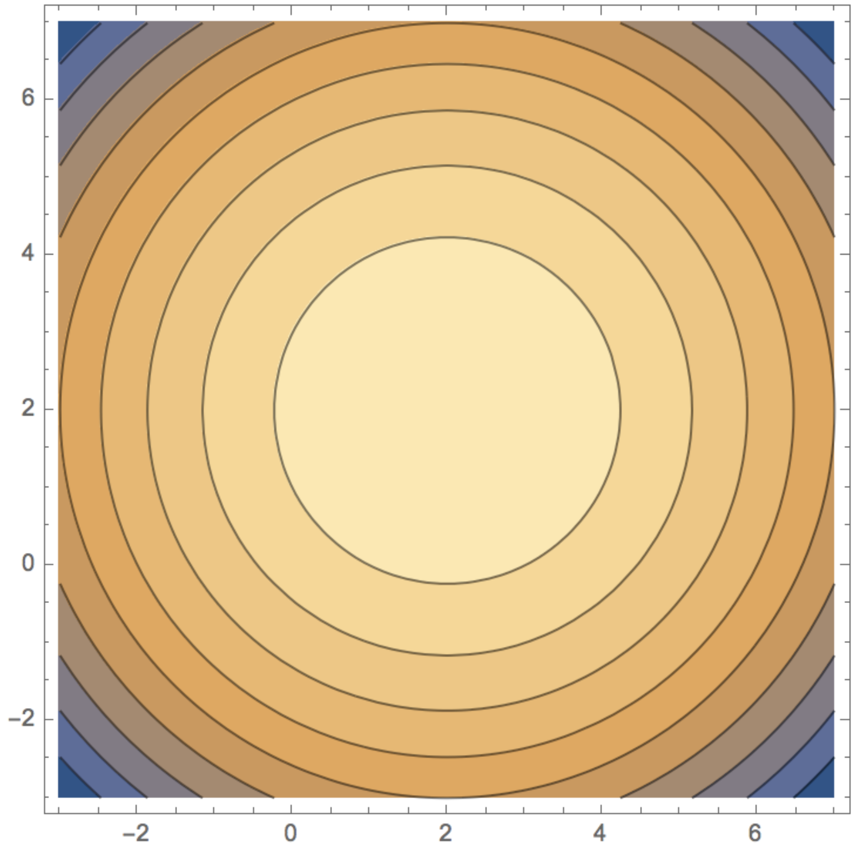
$$\text{s.t.} \quad y_i(\vec{w} \cdot \vec{x}_i + b) \geq 1, \quad i = 1, \cdots, n$$

$$\boxed{\begin{aligned} &\min_{\vec{w},b} \quad \frac{1}{2}\|\vec{w}\|^2 \\ &\text{s.t.} \quad -y_i(\vec{w} \cdot \vec{x}_i + b) + 1 \leq 0, \quad i = 1, \cdots, n \end{aligned}}$$

Lagrange multipliers example 1

$$f(x,y) = 5 - (x-2)^2 - (y-2)^2$$

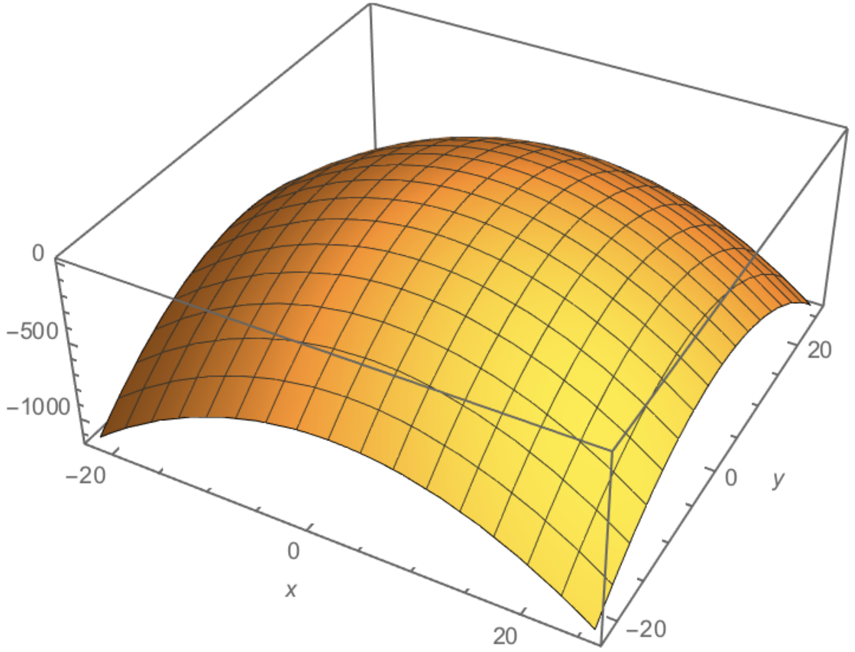


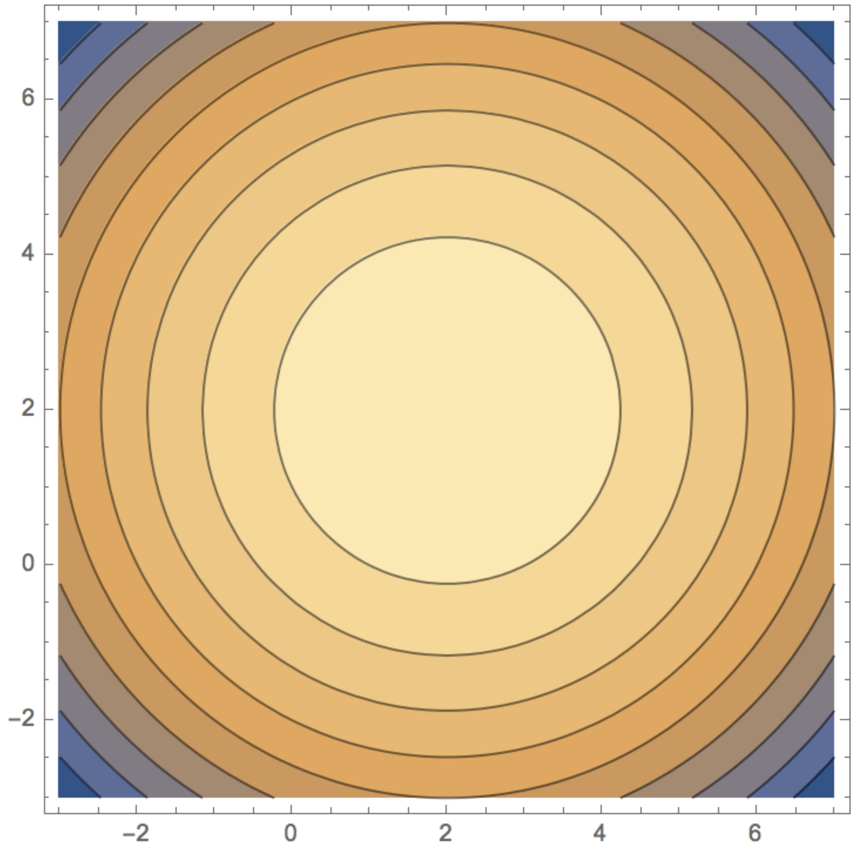


Contour plot of $f(x,y)$

Lagrange multipliers example 1

$$f(x, y) = 5 - (x-2)^2 - (y-2)^2$$





Contour plot of *f*(*x*,*y*)

$$\text{maximize}_{x,y} \quad f(x, y)$$
$$s.t. \quad g(x, y) = 0$$

$$g(x, y) = -5 + x + y$$

Lagrange Multipliers

$$\max(h(\mathrm{x},\mathrm{y},\lambda) = f(x,y) - \lambda g(x,y)$$

We want $\nabla h(x,y,\lambda) = 0$

$$\nabla_{x,y} h(x,y,\lambda) = \nabla f(x,y) - \nabla \lambda g(x,y) = 0$$

$$\nabla f(x,y) = \nabla \lambda g(x,y)$$

$$\frac{\partial h}{\partial \lambda} = g(x,y) = 0$$

$$f(x,y) = 5 - (x-2)^2 - (y-2)^2$$

3 equations

$$\boxed{\begin{array}{c} \text{maximize}_{x,y} \quad f(x,y) \\ \\ s.t. \quad g(x,y) = 0 \end{array}}$$

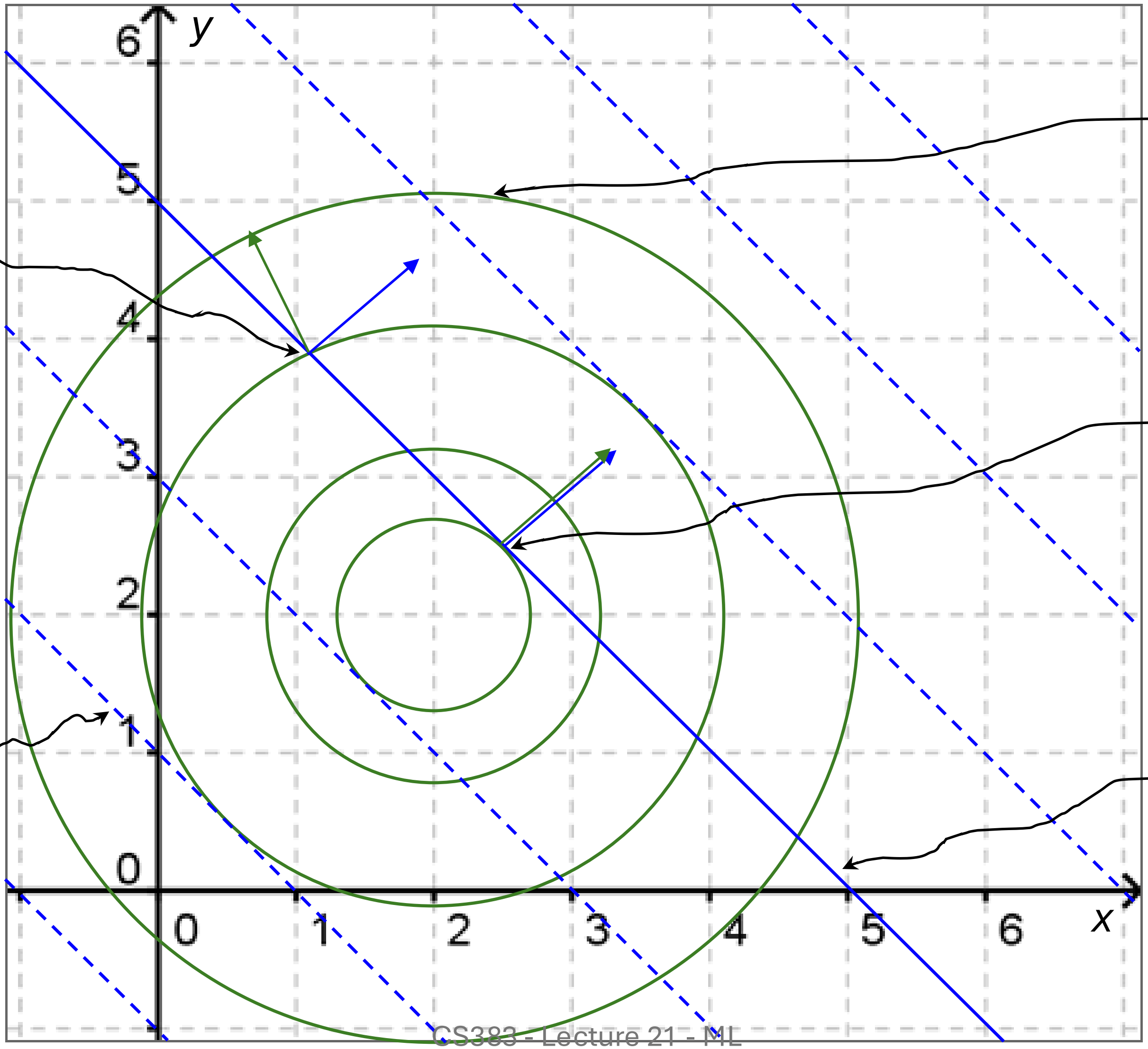
$$\boxed{g(x,y) = -5 + x + y}$$

1. $-5 + x + y = 0$

2. $-2(x-2) = \lambda * 1$

3. $-2(y-2) = \lambda * 1$

Lagrange multipliers example 1



Normals (and derivatives) **not** parallel

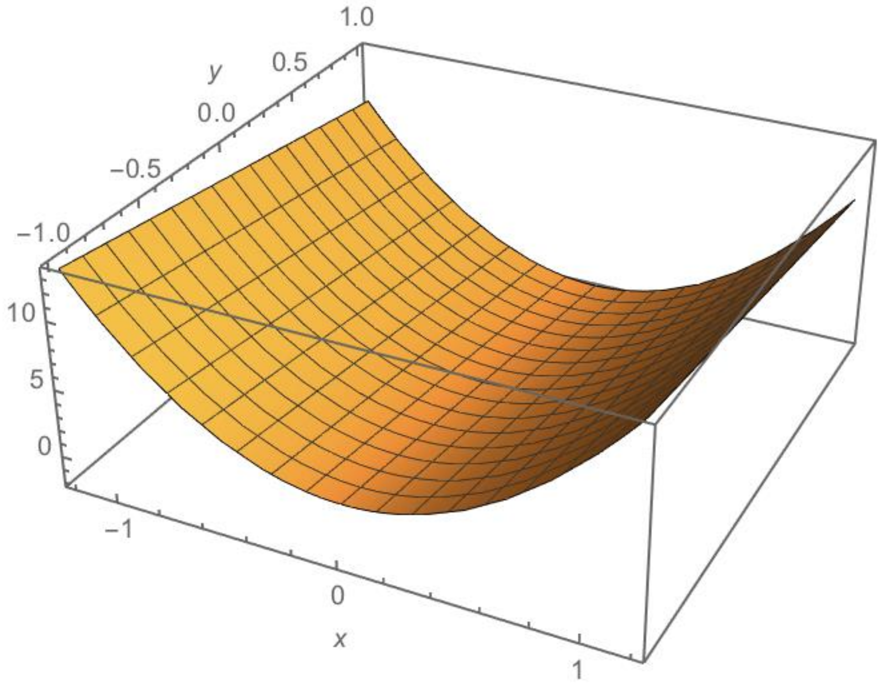
level curves of $f(x,y)$

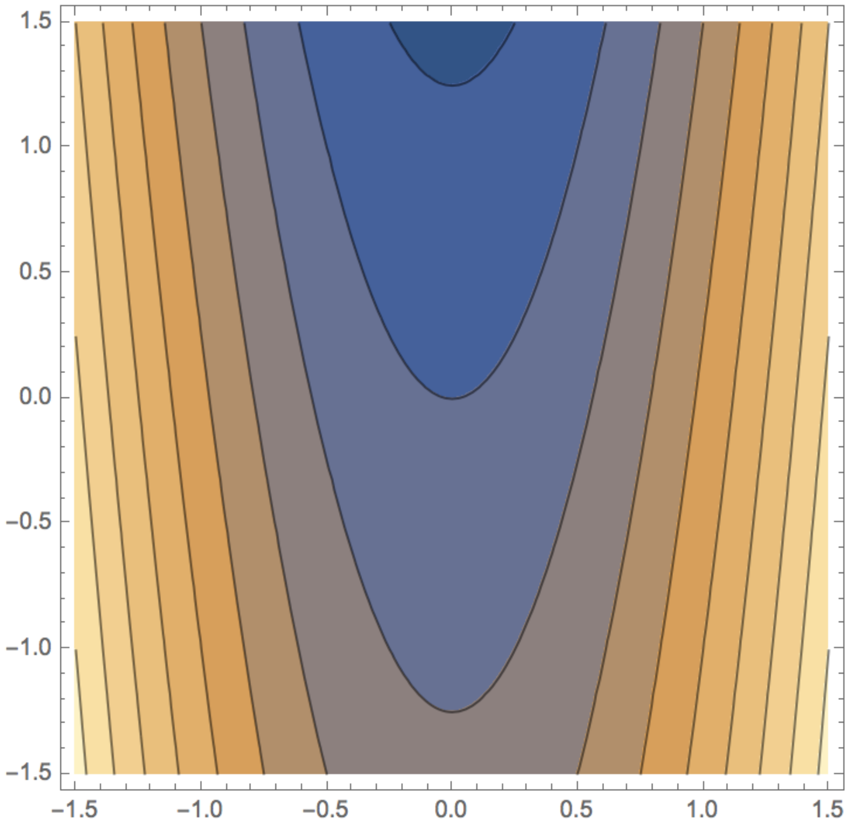
<u>SOLUTION</u>: Normals (and derivatives) parallel
$x = 2.5$
$y = 2.5$

level curves of $g(x,y)$

$g(x,y) = 0$

Lagrange multipliers example 2



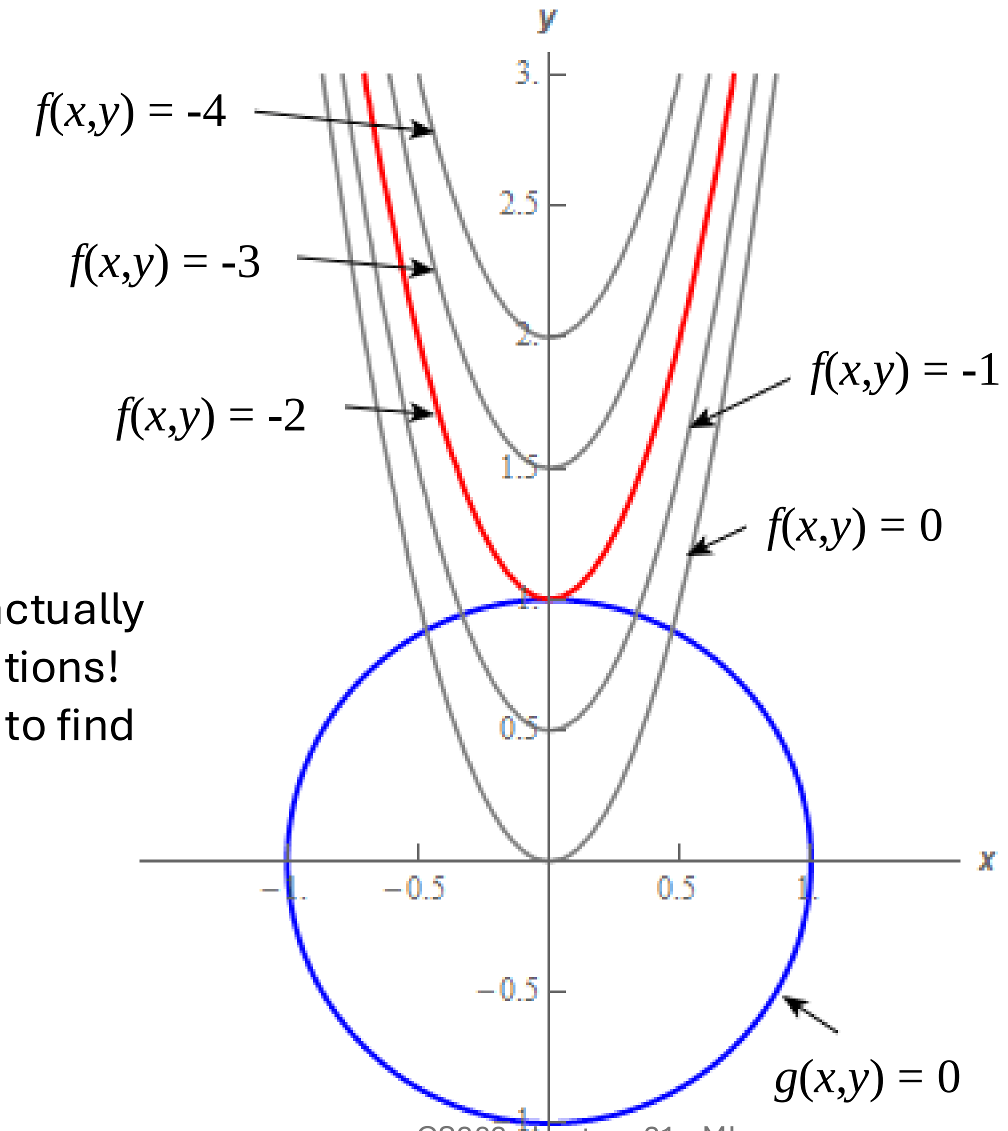


Contour plot of f(x,y)

$$\text{maximize}_{x,y} \quad f(x,y)$$

$$s.t. \quad g(x,y) = 0$$

Lagrange multipliers example 2



Note: there are actually 4 potential solutions! Must plug in to *f* to find max.

Lagrange Multiplier for SVM

Idea: put arbitrary constraint on functional margin

$$\boxed{\hat{\gamma} = 1}$$

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

$$\text{s.t.} \quad y_i(\vec{w} \cdot \vec{x}_i + b) \geq 1, \quad i = 1, \cdots, n$$

Lagrange Multiplier for SVM

$$\min_{\vec{w},b} \quad \frac{1}{2}\|\vec{w}\|^2$$
$$\text{s.t.} \quad y_i(\vec{w}\cdot\vec{x}_i + b) \geq 1, \quad i = 1, \cdots, n$$

$$\max(h(x, y, \lambda) = f(x, y) \; - \lambda g(x, y)$$

Lagrange Multiplier for SVM

$$\min_{\vec{w},b} \quad \frac{1}{2}\|\vec{w}\|^2$$
$$\text{s.t.} \quad y_i(\vec{w}\cdot\vec{x}_i + b) \geq 1, \quad i = 1, \cdots, n$$

$$\max(h(x,y,\lambda) = f(x,y) - \lambda g(x,y))$$

$$\nabla h(x,y,\lambda)$$

Lagrange Multiplier for SVM

$$\min_{\vec{w},b} \quad \frac{1}{2}\|\vec{w}\|^2$$
$$\text{s.t.} \quad y_i(\vec{w}\cdot\vec{x}_i + b) \geq 1, \quad i = 1,\cdots,n$$

$$\max(h(x,y,\lambda) = f(x,y) \; - \lambda g(x,y)$$

$$\nabla h(x,y,\lambda) = \begin{bmatrix} \frac{df}{dx} - \lambda\frac{dg}{dx} \\ \frac{df}{dy} - \lambda\frac{dg}{dy} \\ g(x,y) \end{bmatrix} =$$

Lagrange Multiplier for SVM

$$\min_{\vec{w},b} \quad \frac{1}{2}\|\vec{w}\|^2$$
$$\text{s.t.} \quad y_i(\vec{w}\cdot\vec{x}_i + b) \geq 1, \quad i = 1,\cdots,n$$

$$\max(h(x,y,\lambda) = f(x,y) - \lambda g(x,y))$$

$$\nabla h(x,y,\lambda) = \begin{bmatrix} \frac{df}{dx} - \lambda\frac{dg}{dx} \\ \frac{df}{dy} - \lambda\frac{dg}{dy} \\ g(x,y) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Lagrange Multiplier for SVM

$$\min_{\vec{w},b} \quad \frac{1}{2}\|\vec{w}\|^2$$
$$\text{s.t.} \quad y_i(\vec{w}\cdot\vec{x}_i + b) \geq 1, \quad i = 1,\cdots,n$$

$$h(\vec{w}, b, \vec{\alpha}) = \frac{1}{2}\left||\vec{w}|\right|^2 - \sum_{i=1}^{n} \alpha_i\,[y_i\,\overrightarrow{(w}\ \overrightarrow{x_i})]$$

Lagrange Multiplier for SVM

$$h(\vec{w}, b, \vec{\alpha}) = \frac{1}{2} \left||\vec{w}|\right|^2 - \sum_{i=1}^{n} \alpha_i \left[y_i \, \overrightarrow{(w \;\; \vec{x_i})}\right]$$

$$\nabla_{\vec{w}} h(\vec{w}, b, \vec{\alpha}) = \vec{w} - \sum_{i=1}^{n} \alpha_i \left[y_i \, \vec{x_i}\right] = 0$$

$$\vec{w} = \sum_{i=1}^{n} \alpha_i \left[y_i \, \vec{x_i}\right]$$

Lagrange Multiplier for SVM

$h(\vec{w}, b, \vec{\alpha}) = \frac{1}{2} \left||\vec{w}|\right|^2 - \sum_{i=1}^{n} \alpha_i \left[y_i \overrightarrow{(w} \; \vec{x_i} + b)\right]$

$$\nabla_b h(\vec{w}, b, \vec{\alpha}) = \sum_{i=1}^{n} \alpha_i \, y_i = 0$$

$$\sum_{i:y_i=1}^{n} \alpha_i \, y_i = \sum_{i:y_i=-1}^{n} \alpha_i \, y_i$$

Lagrange Multiplier for SVM

$$h(\vec{w}, b, \vec{\alpha}) = \frac{1}{2}\,\left|\left|\vec{w}\right|\right|^2 \;- \sum_{i=1}^{n} \alpha_i\,[y_i\,\overrightarrow{(w}\;\;\vec{x_i})]$$

$$g(x, y) = 0$$

$$\alpha_i[y_i\,\overrightarrow{(w}\;\;\vec{x_i})] = 0$$

Dual form

$$\max W(\vec{\alpha}) = \sum_{i}^{n} \alpha_i - \frac{1}{2} \sum_{i}^{n} \sum_{j}^{n} y_i y_j \alpha_i \alpha_j \, \overrightarrow{x_i} \, \overrightarrow{x_j}$$

$$s.t. \, \alpha_i > 0 \; \forall i \; \& \sum_{i}^{n} \alpha_i y_i = 0$$

Meta-optimization process

- Incremental SVM optimization algorithm

Meta-optimization process

- Incremental SVM optimization algorithm

- Choose a subset S of examples and run optimization to get alpha values

Meta-optimization process

- Incremental SVM optimization algorithm

- Choose a subset S of examples and run optimization to get alpha values

- Identify which alpha values are 0 => these cannot be support vectors in final solution!

Meta-optimization process

- Incremental SVM optimization algorithm
- Choose a subset S of examples and run optimization to get alpha values
- Identify which alpha values are 0 => these cannot be support vectors in final solution!
- Discard these points and add new ones; repeat