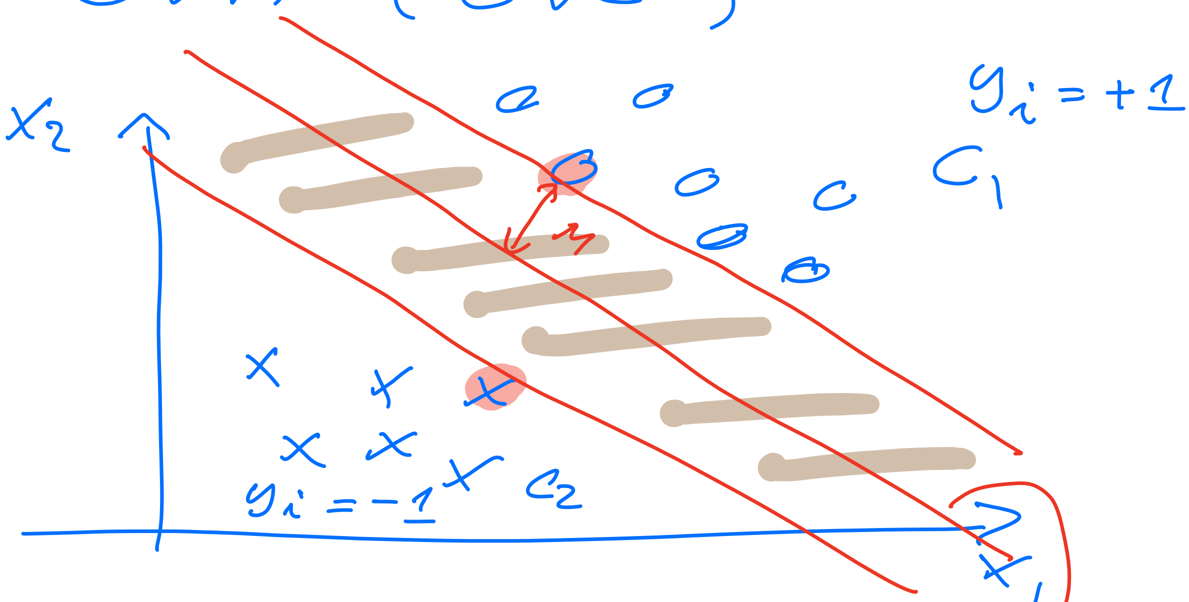
FYS-STK4155, Nov 17, 2022

SVM (SVC)



$y_i \in \{-1, 1\}$

$\uparrow$ C_2 $\quad$ $\uparrow$ C_1

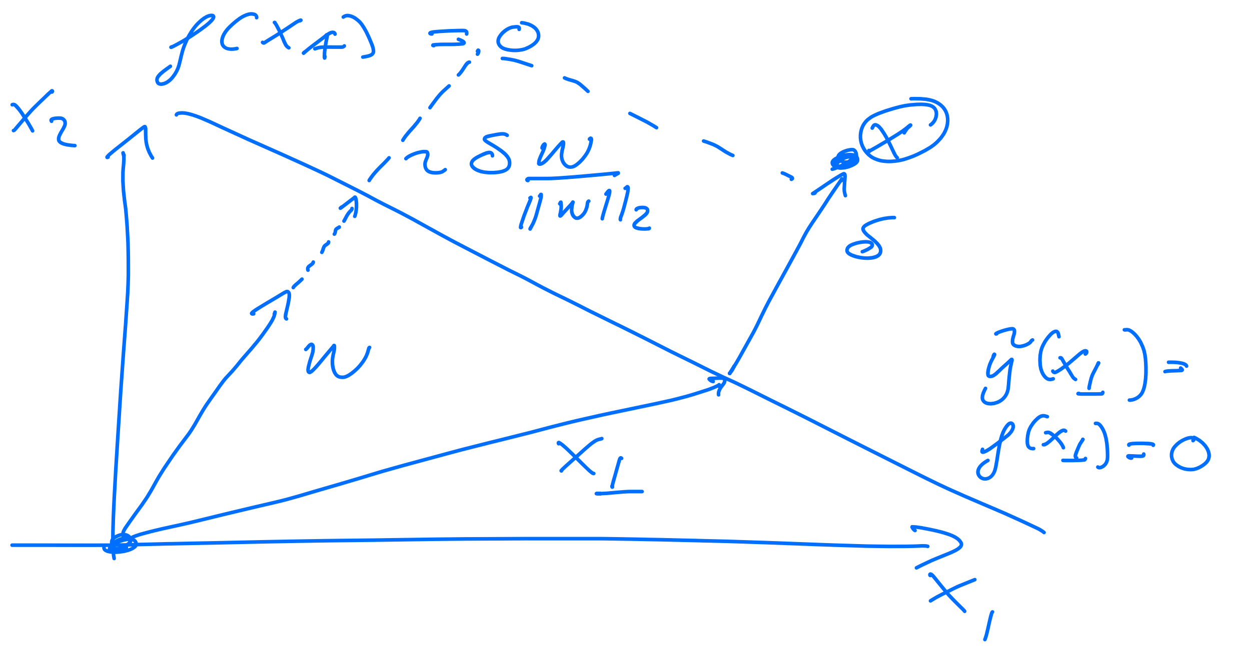
$\tilde{y}_i = f(x_i) = 0$

$y_i f(x_i) > 0$ correct prediction

$$\tilde{y}_i = f(x) = w^T x + w_0$$

$$= w_1 x_1 + w_2 x_2 + w_0$$

A point on the midpoint line x_A



$$x = x_\perp + \delta \frac{w}{\|w\|_2}$$

$$\tilde{y}(x_\perp) = 0 = w^T x_\perp + w_0$$

$$\delta = \frac{\tilde{y}(x)}{\|w\|_2}$$

$$\frac{y_i (w^T x_i + w_0)}{\|w\|_2} \geq M$$

$$y_i \tilde{y}_i = y_i (w^T x_i + w_0) \geq M \|w\|_2$$

for all $i = 0, 1, \ldots n-1$

maximize M subject to the constraint

$$M \|w\|_2 = \underline{1}$$

$$M = \frac{1}{\|w\|_2}$$

$$\max_{w, w_0} M$$

with constraint

$$y_i (w^T x_i + w_0) \geq \underline{1}$$

or

$$\min_{w, w_0} \frac{1}{2} w^T w$$

with constraint

$$y_i (w^T x_i + w_0) \geq \underline{1}$$

optimization with constraints.

Example

find extreme values of

$f(x_1, x_2) = x_1^2 - x_2^2$ when

$x_1 \geq 0 \quad \wedge \quad x_2 \geq 0$

$f(x) \in \mathbb{R}^n$ and

constraints of the type

$g_1(x) = b_1, \; g_2(x) = b_2 \ldots$

$g_m(x) = b_m$

b_i are constants.

introduce the Lagrangian

$$\mathcal{L}(x, \lambda) = f(x) - \lambda_1 (g_1(x) - b_1)$$

$$- \ldots - \lambda_m (g_m(x) - b_m)$$

$\lambda_1, \lambda_2, \ldots \lambda_m$ are called

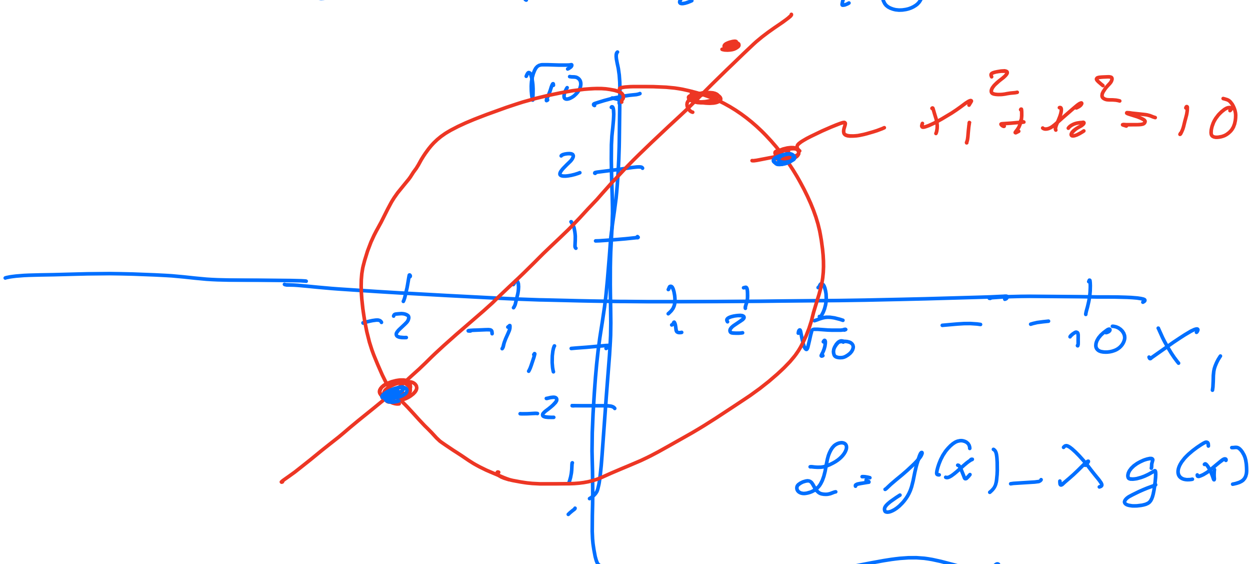
Lagrange multipliers.

$$\frac{\partial \mathcal{L}}{\partial x_i} = 0 \quad \text{for } i = 0, 1, \ldots n-1$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_j} = 0 \quad \text{for } j = 1, 2, \ldots m$$

Example

$f(x_1, x_2) = x_1 + 3x_2$ subject to $x_1^2 + x_2^2 = 10$



$$\mathcal{L} = f(x) - \lambda g(x)$$

$$\frac{\partial \mathcal{L}}{\partial x_1} = \frac{\partial}{\partial x_1}\left(x_1 + 3x_2 - \lambda(x_1^2 + x_2^2 - 10)\right)$$

$$= 0 \Rightarrow 1 - 2\lambda x_1 = 0$$

$$x_1 = 1/2\lambda$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = 0 \Rightarrow x_2 = \frac{3}{2\lambda}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = 0 \Rightarrow x_1^2 + x_2^2 = 10$$

$$\frac{1}{4\lambda^2} + \frac{9}{4\lambda^2} = 10$$

$$\lambda = \pm 1/2$$

(i) $\lambda = +1/2 \quad x_1 = 1 \quad \cancel{y_2}\, x_2 = 3$

$\lambda = -1/2 \quad x_1 = -1 \quad x_2 = -3$

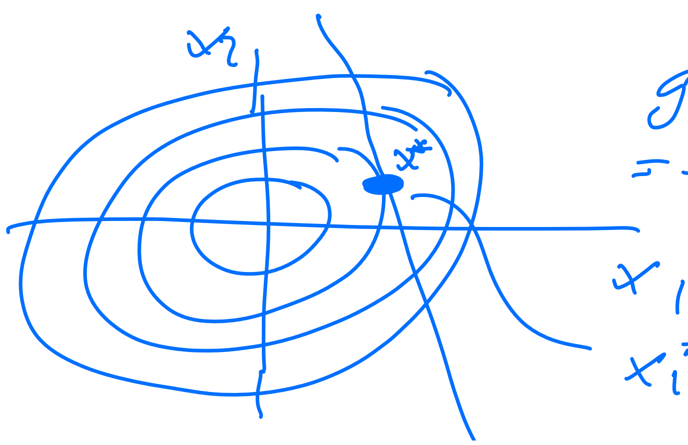
K (Kushen) K (Kuhn) T (Tucker) conditions

constrained solutions

(i) $g(x) > 0$ (constraint is inactive)

(ii) $g(x) = 0$ constraint is active

$$f(x_1, x_2) = 1 - x_1^2 - x_2^2$$



$g(x_1, x_2) = x_1 + x_2 - 1$

$x_1^* \; x_2^*$

* when $g(x) > 0$, it plays no role in the optimization and the stationary points are given by $\nabla_x f(x) = 0$ with $\lambda = 0$

* when solution is on the boundary, $g(x) = 0$

$\lambda \neq 0$

For either case

$$\lambda \cdot g(x) = 0$$

if maximize $f(x)$ subject to $g(x) \geq 0$, we optimize $\mathcal{L}$ wrt x and λ

subject to

$$\boxed{\begin{array}{ll} g(x) \geq 0 & \lambda \geq 0 \\ \lambda g(x) = 0 & \end{array}}$$

KKT - conditions

For minimization

$\mathcal{L}(x, \lambda) = f(x) - \lambda g(x)$

with same KKT conditions

SVM with classification and margin M

$$\min_{w, w_0} \frac{1}{2} w^T w$$

subject to

$$y_i (w^T x_i + w_0) = 1 \quad \forall i = 0, 1, \dots m-1$$

$$\mathcal{L} = \frac{1}{2} w^T w - \sum_{i=0}^{m-1} \lambda_i \left[y_i (w^T x_i + w_0) - 1 \right]$$

$$\frac{\partial \mathcal{L}}{\partial w_0} = 0 = -\sum y_i \lambda_i$$

$$\frac{\partial \mathcal{L}}{\partial w^T} = 0 = w - \sum_i \lambda_i y_i x_i$$

$\uparrow$ vector

$$\frac{\partial \mathcal{L}}{\partial \lambda_i} = 0 = y_i(w^T x_i + w_0) - 1$$

$$w = \sum_i \lambda_i y_i x_i$$

insert w in $\mathcal{L}$

$$\mathcal{L} = \sum_{i=0}^{n-1} \lambda_i - \frac{1}{2} \sum_{i,j} \lambda_i \lambda_j y_i y_j x_i^T x_j$$

subject to (KKT)

$\boxed{\lambda_i \geq 0}$ KKT

$\sum \lambda_i y_i = 0$ and

$$\boxed{\begin{array}{l} \lambda_i [y_i(w^T x_i + w_0) - 1] \\ \lambda_i > 0 \quad \text{then} \\ \qquad y_i(w^T x_i + w_0) = 1 \end{array}}$$ KKT

x_n is on the boundary (Margin M) and define the support vector

This is referred to as a hard margin problem?

Define

$$\lambda^T = [\lambda_0, \lambda_1 \ldots \lambda_{n-1}]$$

$$y^T = [y_0\ y_1 \ldots y_{n-1}]$$

Rewrite

$$\mathcal{L} = \mathbb{1}\lambda - \frac{1}{2}\lambda^T \begin{bmatrix} y_0 y_0 x_0^T x_0 & y_0 y_1 x_0^T x_1 & \ldots & y_0 y_{n-1} x_0^T x_{n-1} \\ \vdots & \times \quad \times & & \times \\ \vdots & \times \quad \times & & \times \\ y_{n-1} y_0 x_{n-1}^T x_0 & \ldots & \ldots & y_{n-1} y_{n-1} x_{n-1}^T x_{n-1} \end{bmatrix}$$

$x \quad \lambda$

subject to $\quad y^T \lambda = 0$

\+ KKT conditions

We can find the "intercept"

$$y_i (w^T x_i + w_0) = 1$$

$$w_0 = \frac{1}{y_i} - w^T x_i$$