Question 1: Lagrangian

Define $F = (f(x_1), ..., f(x_N))$ and $Y = (y_1, ..., y_N)$.
Let f in $\mathcal{H}$, ξ, α, μ be vectors in $\mathbb{R}^N$ with $\alpha \geq 0$ and $\eta \geq 0$.
The Lagrangian is defined

$$\begin{aligned} L(f, b, \xi, \alpha, \eta) = & \frac{1}{2}\|f\|^2 + C\xi^\top \mathbb{1} - (Y \odot \alpha)^\top F \\ & - \alpha^\top (bY - \mathbb{1} + \xi) - \mu^\top \xi. \\ = & \frac{1}{2}\|f\|^2 - \langle f, (Y \odot \alpha)^\top \tilde{K} \rangle_{\mathcal{H}} - b\alpha^\top Y + \alpha^\top \mathbb{1} \\ & + \xi^\top (C\mathbb{1} - \alpha - \mu). \end{aligned}$$

We used that $(Y \odot \alpha)^\top F = \langle f, (Y \odot \alpha)^\top \tilde{K}$, with

$$(Y \odot \alpha)^\top \tilde{K} = \sum_{i=1}^{N} \alpha_i y_i k(x_i, .).$$

Question 1: Dual problem

$$L(f, b, \xi, \alpha, \eta) = \frac{1}{2}\|f\|^2 - \langle f, (Y \odot \alpha)^\top \tilde{K} \rangle_{\mathcal{H}} - b\alpha^\top Y + \alpha^\top \mathbb{1} \\ + \xi^\top (C\mathbb{1} - \alpha - \mu).$$

with $(Y \odot \alpha)^\top \tilde{K} = \sum_{i=1}^{N} \alpha_i y_i k(x_i, .)$.
Stationarity conditions in primal variables:

$$\nabla_f L = 0, \quad \nabla_b L = 0, \quad \nabla_\xi L = 0.$$

Question 1: Dual problem

$$L(f, b, \xi, \alpha, \eta) = \frac{1}{2}\|f\|^2 - \langle f, (Y \odot \alpha)^\top \tilde{K} \rangle_{\mathcal{H}} - b\alpha^\top Y + \alpha^\top \mathbb{1} \\ + \xi^\top (C\mathbb{1} - \alpha - \mu).$$

with $(Y \odot \alpha)^\top \tilde{K} = \sum_{i=1}^{N} \alpha_i y_i k(x_i, .)$.
Stationarity conditions in primal variables:

$$f = (Y \odot \alpha)^\top \tilde{K}, \qquad \alpha^\top Y = 0, \qquad C\mathbb{1} - \alpha - \mu = 0.$$

Question 1: Dual problem

$$L(f,b,\xi,\alpha,\eta) = \frac{1}{2}\|f\|^2 - \langle f, (Y \odot \alpha)^\top \tilde{K} \rangle_{\mathcal{H}} - b\alpha^\top Y + \alpha^\top \mathbb{1}$$
$$+ \xi^\top (C\mathbb{1} - \alpha - \mu).$$

with $(Y \odot \alpha)^\top \tilde{K} = \sum_{i=1}^N \alpha_i y_i k(x_i, .)$.
Stationarity conditions in primal variables:

$$f = (Y \odot \alpha)^\top \tilde{K}, \qquad \alpha^\top Y = 0, \qquad C\mathbb{1} - \alpha - \mu = 0.$$

Replacing the above in the Lagrangian:

$$L(f,b,\xi,\alpha,\eta) = -\frac{1}{2}\|f\|^2 + \alpha^\top \mathbb{1} = -\frac{1}{2}\alpha^\top G \alpha + \alpha^\top \mathbb{1}$$

with $G_{ij} = y_i y_j k(x_i, x_j)$. Recall that: $\alpha \geq 0$ and $C\mathbb{1} - \alpha = \mu \geq 0$, hence the dual problem is

$$\min_{\alpha} \frac{1}{2}\alpha^\top G \alpha - \alpha^\top \mathbb{1},$$
$$s.t.: \alpha^\top Y = 0, \qquad 0 \leq \alpha \leq C\mathbb{1}.$$

Question 1: Boundary points

- Complementary slackness:

$$\alpha_i(y_i(f(x_i) + b) - 1 + \xi_i) = 0, \quad \mu_i \xi_i = 0$$

- In the homework, we are interested in finding sufficient conditions on the dual parameter α_i so that (x_i, y_i) are support vector points that fall on the margin of the separating hyper-surface:

$$y_i(f(x_i) + b) - 1 = 0$$

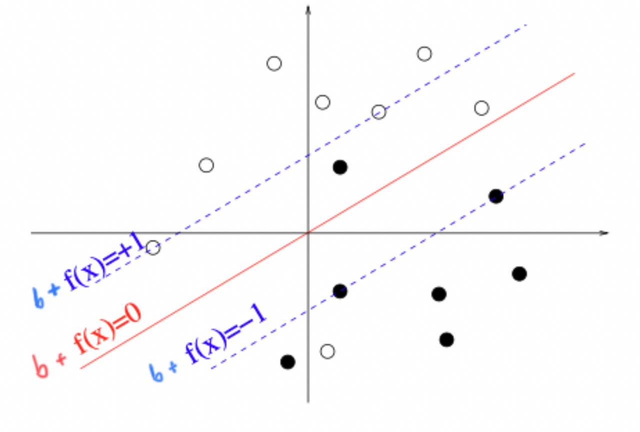
.

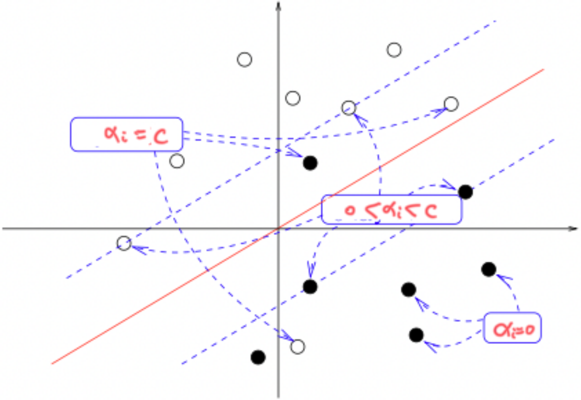
- By looking for indices i such that $\alpha_i > 0$, the complementary slackness implies:

$$y_i(f(x_i) + b) - 1 + \xi_i = 0$$

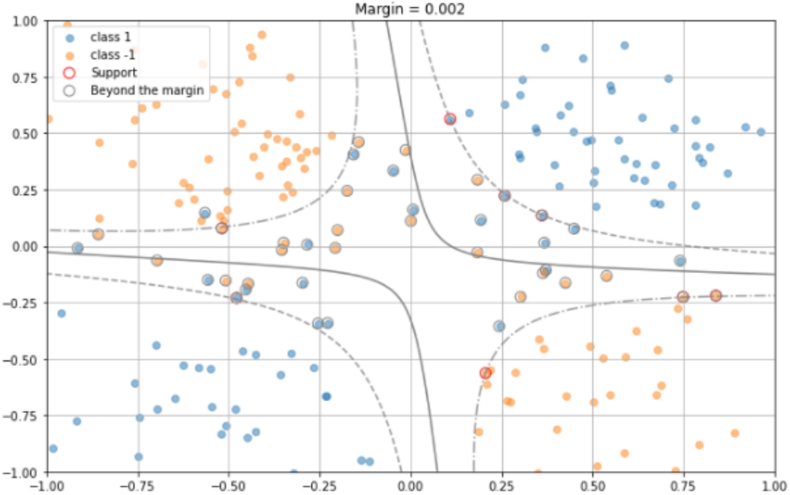
- If in addition $\mu_i > 0$, then again by complementary slackness it holds that $\xi_i = 0$, so that $y_i(f(x_i) + b) - 1$.
- The constraint $C - \alpha_i = \mu_i > 0$ implies that $\alpha_i < 0$.
- Hence, we are looking for points such that $0 < \alpha_i < C$.
- For a strict margin point: $b = y_i - f(x_i)$.

Question 1: Boundary points





- Points for which $\alpha_i > 0$ are on the wrong side of the margin.
- Points for which $\alpha_i = C$ are strictly on the wrong side of the margin.
- Points for which $0 < \alpha_i < C$ are exactly on the margin.



Margin = 0.002
class 1
class -1
Support
Beyond the margin
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00

Question 2: Lagrangian

Define $F = (f(x_1), ..., f(x_N))$ and $Y = (y_1, ..., y_N)$.
Let f in $\mathcal{H}$, ξ=(ξ^+, ξ^-), $\alpha = (\alpha^+, \alpha^-)$, $\mu = (\mu^+, \mu^-)$ be vectors in $\mathbb{R}^{2N}$ with $\alpha \geq 0$ and $\mu \geq 0$. The Lagrangian is defined

$$\begin{aligned} L(f, b, \xi, \alpha, \mu) =& \frac{1}{2}\|f\|^2 + C(\xi^+ + \xi^-)^\top \mathbb{1} \\ &+ (Y - F - (b+\eta)\mathbb{1} - \xi^+)^\top \alpha^+ - (\mu^+)^\top \xi^+ \\ &+ (-Y + F + (b-\eta)\mathbb{1} - \xi^-)^\top \alpha^- - (\mu^-)^\top \xi^- \end{aligned}$$

Stationarity conditions

$$f = (\alpha^+ - \alpha^-)^\top \tilde{K}, \qquad \mathbb{1}^\top (\alpha^+ - \alpha^-) = 0,$$
$$C\mathbb{1} - \alpha^+ - \mu^+ = C\mathbb{1} - \alpha^- - \mu^- = 0.$$

$$L(f, b, \xi, \alpha, \mu) = -\frac{1}{2}\|f\|^2 + Y^\top (\alpha^+ - \alpha^-) - \eta \mathbb{1}^\top (\alpha^+ + \alpha^-).$$

Question 2: Lagrangian

Stationarity conditions

$$f = (\alpha^+ - \alpha^-)^\top \tilde{K}, \qquad \mathbb{1}^\top(\alpha^+ - \alpha^-) = 0,$$
$$C\mathbb{1} - \alpha^+ - \mu^+ = C\mathbb{1} - \alpha^- - \mu^- = 0.$$

Hence, Lagrangian becomes:

$$\begin{aligned} L(f, b, \xi, \alpha, \mu) &= -\frac{1}{2}\|f\|^2 + Y^\top(\alpha^+ - \alpha^-) - \eta\mathbb{1}^\top(\alpha^+ + \alpha^-) \\ &= -\frac{1}{2}\delta^\top K\delta + Y^\top\delta - \eta\mathbb{1}^\top(\alpha^+ + \alpha^-), \end{aligned}$$

with $\delta{=}\alpha^+ - \alpha^-$.
Dual problem:

$$\begin{aligned} &\min_\alpha \frac{1}{2}\delta^\top K\delta - Y^\top\delta + \eta\mathbb{1}^\top(\alpha^+ + \alpha^-), \\ s.t. \qquad &0 \le \alpha^+ \le C\mathbb{1}, 0 \le \alpha^- \le C\mathbb{1}, \delta^\top\mathbb{1}. \end{aligned}$$

Question 2: Boundary points

- Complementary slackness (CS):

$$\alpha_i^+(y_i - f(x_i) - b - \eta - \xi_i^+) = 0, \qquad \mu_i^+ \xi_i^+ = 0$$
$$\alpha_i^-(-y_i + f(x_i) + b - \eta - \xi_i^-) = 0 \qquad \mu_i^- \xi_i^-$$

- In the homework, we are interested in finding sufficient conditions on the dual parameters α_i^+, α_i^- so that $y_i - f(x_i) - b{=}\eta$ or $-y_i + f(x_i) + b{=}\eta$.
- First, if $\alpha_i^+ > 0$ or $\alpha_i^- > 0$, then by CS:

$$\begin{cases} y_i - f(x_i) - b - \eta - \xi_i^+ = 0, & \alpha_i^+ > 0 \\ -y_i + f(x_i) + b - \eta - \xi_i^- = 0, & \alpha_i^- > 0 \end{cases} \tag{1}$$

- Moreover, if $\mu_i^+{>}0$ in case 1 or $\mu_i^-{>}0$ in case 2, then by CS: $C - \alpha_i^+{=}\mu_i^+ > 0$ or $C - \alpha_i^-{=}\mu_i^- > 0$.
- Need points such that $0 < \alpha_i^+ < C$ or $0 < \alpha_i^- < C$.
- Also note that if $0 < \alpha_i^+ < C$, then $\alpha_i^- = 0$ and vice-versa (otherwise, we get $-2\eta = \xi_i^- < 0$ contradiction). Hence, need points s.t. $0 < |\alpha_i^+ - \alpha_i^-| < C$.

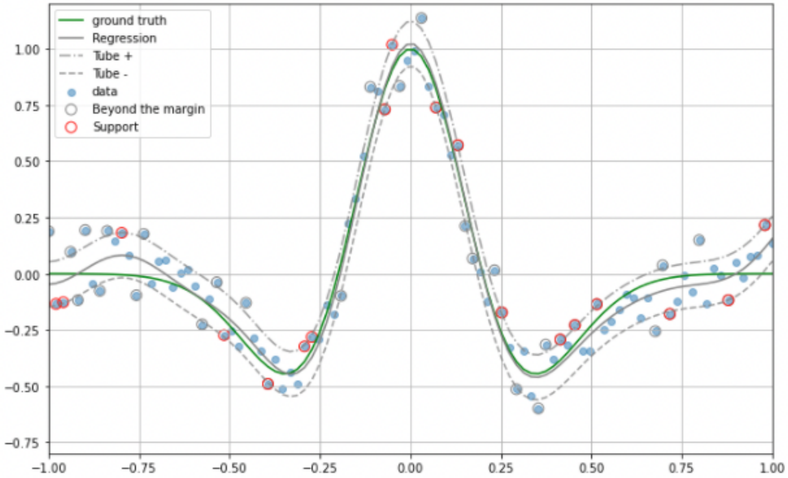
Question 2: Boundary points

- For a boundary point:

$$\begin{cases} b = y_i - f(x_i) - \eta, & 0 < \alpha_i^+ < C \\ b = \eta + y_i - f(x_i), & 0 < \alpha_i^- < C. \end{cases}$$

- Regression function:

$$f(x) = \sum_{\{i|\alpha_i^+ > 0\} \cup \{i|\alpha_i^- > 0\}} (\alpha_i^+ - \alpha_i^-) k(x_i, x)$$


ground truth
Regression
Tube +
Tube -
data
Beyond the margin
Support
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00