Vapnik's ε-regression Classic formulation.

$$\min_{w,b,e,e^*} \mathcal{J}(w,e,e^*) = \frac{1}{2}w^Tw + c\sum_{k=1}^{N}(e_k+e_k^*)$$

$$\text{s.t.} \quad \begin{aligned} & y_k - w^T\varphi(x_k) - b \le \varepsilon + e_k \\ & w^T\varphi(x_k) + b - y_k \le \varepsilon + e_k^* \\ & e_k, e_k^* \geq 0 \end{aligned} \qquad k = 1,\dots,N$$

Lagrangian

$$L = \frac{1}{2}w^Tw + c\sum_{k=1}^{N}(e_k+e_k^*) - \sum_{k=1}^{N}\alpha_k(\varepsilon - e_k - y_k + w^T\varphi(x_k) + b) - \sum_{k=1}^{N}\alpha_k^*(\varepsilon + e_k^* + y_k - w^T\varphi(x_k) - b) - \sum_{k=1}^{N}\eta_k e_k - \sum_{k=1}^{N}\eta_k^* e_k^*$$

Saddle point:

$$\frac{\partial L}{\partial w} = w - \sum_{k=1}^{N}\alpha_k\varphi(x_k) + \sum_{k=1}^{N}\alpha_k^*\varphi(x_k) = 0 \Rightarrow w = \sum_{k=1}^{N}(\alpha_k - \alpha_k^*)\varphi(x_k)$$

$$\frac{\partial L}{\partial b} = -\sum_{k=1}^{N}\alpha_k + \sum_{k=1}^{N}\alpha_k^* = 0 \Rightarrow \sum_{k=1}^{N}(\alpha_k - \alpha_k^*) = 0$$

$$\frac{\partial L}{\partial e_k} = c - \alpha_k - \eta_k = 0, \qquad \frac{\partial L}{\partial e_k^*} = c - \alpha_k^* - \eta_k^* = 0$$

Then, replacing in the Lagrangian:

$$D = \frac{1}{2}\sum_{k=1}^{N}(\alpha_k-\alpha_k^*)\varphi(x_k)^T\sum_{l=1}^{N}(\alpha_l-\alpha_l^*)\varphi(x_l) + c\sum_{k=1}^{N}(e_k+e_k^*) - \sum_{k=1}^{N}\alpha_k(\varepsilon + e_k - y_k + w^T\varphi(x_k))$$
$$- \sum_{k=1}^{N}\alpha_k b - \sum_{k=1}^{N}\alpha_k^*(\varepsilon + e_k^* + y_k - w^T\varphi(x_k)) + \sum_{k=1}^{N}\alpha_k^* b - \sum_{k=1}^{N}(c-\alpha_k)e_k - \sum_{k=1}^{N}(c-\alpha_k^*)e_k^*$$

$$= \frac{1}{2}\sum_{k,l=1}^{N}(\alpha_k-\alpha_k^*)(\alpha_l-\alpha_l^*)\varphi(x_k)^T\varphi(x_l) - \sum_{k=1}^{N}\alpha_k(\varepsilon + e_k - y_k + w^T\varphi(x_k))$$
$$- \sum_{k=1}^{N}\alpha_k^*(\varepsilon + e_k^* + y_k - w^T\varphi(x_k)) + \sum_{k=1}^{N}\alpha_k e_k + \sum_{k=1}^{N}\alpha_k^* e_k^*$$

$$= \frac{1}{2}\sum_{k,l}^{N}(\alpha_k-\alpha_k^*)(\alpha_l-\alpha_l^*)\varphi(x_k)^T\varphi(x_l) - \varepsilon\sum_{k=1}^{N}(\alpha_k+\alpha_k^*) + \sum_{k=1}^{N}y_k(\alpha_k-\alpha_k^*) - \sum_{k=1}^{N}(\alpha_k-\alpha_k^*)w^T\varphi(x_k)$$

$$= -\frac{1}{2}\sum_{k=1}^{N}(\alpha_k-\alpha_k^*)(\alpha_l-\alpha_l^*)\varphi(x_k)^T\varphi(x_l) - \varepsilon\sum_{k=1}^{N}(\alpha_k+\alpha_k^*) + \sum_{k=1}^{N}y_k(\alpha_k-\alpha_k^*)$$

MAPE ε-SVM

Primal formulation

$$\min_{w,b,e,e^*} \mathcal{P}(w,e,e^*) = \frac{1}{2} w^T w + c \sum_{k=1}^{N} (e_k + e_k^*)$$

$$\text{s.t.} \quad \frac{y_k - w^T \varphi(x_k) - b}{y_k} \leq \varepsilon + e_k \qquad k = 1, \ldots, N$$

$$\frac{w^T \varphi(x_k) + b - y_k}{y_k} \leq \varepsilon + e_k^*$$

$$e_k, e_k^* \geq 0$$

Lagrangian

$$\mathcal{L} = \frac{1}{2} w^T w + c \sum_{k=1}^{N} (e_k + e_k^*) - \sum_{k=1}^{N} \alpha_k \left((\varepsilon + e_k) y_k - y_k + w^T \varphi(x_k) + b \right)$$
$$- \sum_{k=1}^{N} \alpha_k^* \left((\varepsilon + e_k) y_k + y_k - w^T \varphi(x_k) - b \right) - \sum_{k=1}^{N} (\eta_k e_k + \eta_k^* e_k^*)$$

Saddle point

$$\frac{\partial \mathcal{L}}{\partial w} = w - \sum_{k=1}^{N} \alpha_k \varphi(x_k) + \sum_{k=1}^{N} \alpha_k^* \varphi(x_k) = 0 \Rightarrow w = \sum_{k=1}^{N} (\alpha_k - \alpha_k^*) \varphi(x_k)$$

$$\frac{\partial \mathcal{L}}{\partial b} = -\sum_{k=1}^{N} \alpha_k + \sum_{k=1}^{N} \alpha_k^* = 0 \Rightarrow \sum_{k=1}^{N} (\alpha_k - \alpha_k^*) = 0$$

$$\frac{\partial \mathcal{L}}{e_k} = c - \alpha_k y_k - \eta_k = 0, \qquad \frac{\partial \mathcal{L}}{\partial e_k^*} = c - \alpha_k^* y_k - \eta_k^* = 0$$

$$\mathcal{D} = \frac{1}{2} w^T w + c \sum_{k=1}^{N} (e_k + e_k^*) - \sum_{k=1}^{N} \alpha_k \left((\varepsilon + e_k) y_k - y_k + w^T \varphi(x_k) \right) - \sum_{k=1}^{N} \alpha_k^* \left((\varepsilon + e_k^*) y_k + y_k - w^T \varphi(x_k) \right)$$
$$- \sum_{k=1}^{N} \left[(c - \alpha_k y_k) e_k + (c - \alpha_k^* y_k) e_k^* \right]$$
$$= \frac{1}{2} w^T w - \sum_{k=1}^{N} \alpha_k \left((\varepsilon + e_k) y_k - y_k + w^T \varphi(x_k) \right) - \sum_{k=1}^{N} \alpha_k^* \left((\varepsilon + e_k^*) y_k + y_k - w^T \varphi(x_k) \right)$$
$$+ \sum_{k=1}^{N} \left(\alpha_k y_k e_k + \alpha_k^* y_k e_k^* \right)$$

$$= \frac{1}{2} w^T w - \sum_{k=1}^{N} \alpha_k \left((\varepsilon - 1) y_k + w^T \varphi(x_k) \right) - \sum_{k=1}^{N} \alpha_k^* \left((\varepsilon + 1) y_k - w^T \varphi(x_k) \right)$$
$$= \frac{1}{2} w^T w - w^T w - \sum_{k=1}^{N} \left[\alpha_k (\varepsilon - 1) y_k + \alpha_k^* (\varepsilon + 1) y_k \right]$$

$$= -\frac{1}{2}\sum_{k,l=1}^{N}(\alpha_k-\alpha_k^*)(\alpha_k-\alpha_l^*)\varphi(x_k)^T\varphi(x_l) - \sum_{k=1}^{N}\left(\alpha_k(\varepsilon-1)+\alpha_k^*(\varepsilon+1)\right)y_k$$

$$= -\frac{1}{2}\sum_{k,l=1}^{N}(\alpha_k-\alpha_k^*)(\alpha_l-\alpha_l^*)\varphi(x_k)^T\varphi(x_l) - \sum_{k=1}^{N}\left(\varepsilon(\alpha_k+\alpha_k)-(\alpha_k-\alpha_k^*)\right)y_k$$

$$= -\frac{1}{2}\sum_{k,l=1}^{N}(\alpha_k-\alpha_k^*)(\alpha_l-\alpha_l^*)\varphi(x_k)^T\varphi(x_l) - \varepsilon\sum_{k=1}^{N}(\alpha_k+\alpha_k^*)y_k + \sum_{k=1}^{N}(\alpha_k-\alpha_k^*)y_k$$

$$\max_{\alpha_k,\alpha_k^*} D(\alpha,\alpha^*) = -\frac{1}{2}\sum_{k,l=1}^{N}(\alpha_k-\alpha_k^*)(\alpha_l-\alpha_l^*)\varphi(x_k)^T\varphi(x_l) - \varepsilon\sum_{k=1}^{N}y_k(\alpha_k+\alpha_k^*) + \sum_{k=1}^{N}y_k(\alpha_k-\alpha_k^*)$$

$$\text{s.t.} \quad \sum_{k=1}^{N}(\alpha_k-\alpha_k^*)=0$$

$$\alpha_k,\alpha_k^* \in [0, c/y_k]$$

The KKT complementary conditions are still missing.

ν - Formulation - Original problem → Classic formulation

$$\min_{w,b,e,e^*} \ \frac{1}{2}w^Tw + c\left(\nu\varepsilon + \sum_{k=1}^{N}(e_k + e_k^*)\right), \quad \nu\varepsilon \text{ is a new term}$$

$$\text{s.t.} \quad y_k - w^T\varphi(x_k) - b \leq \varepsilon + e_k$$
$$w^T\varphi(x_k) + b - y_k \leq \varepsilon + e_k^*$$
$$e_k, e_k^* \geq 0$$

Lagrangian:

$$\mathcal{L} = \frac{1}{2}w^Tw + c\left(\nu\varepsilon + \sum_{k=1}^{N}(e_k + e_k^*)\right) - \sum_{k=1}^{N}\alpha_k\left(\varepsilon + e_k + w^T\varphi(x_k) + b - y_k\right)$$
$$- \sum_{k=1}^{N}\alpha_k^*\left(\varepsilon + e_k^* - w^T\varphi(x_k) - b + y_k\right) - \sum_{k=1}^{N}\left(\eta_k e_k + \eta_k^* e_k^*\right)$$

Saddle point:

$$\frac{\partial \mathcal{L}}{\partial w} = w - \sum_{k=1}^{N}\alpha_k\varphi(x_k) + \sum_{k=1}^{N}\alpha_k^*\varphi(x_k) = 0 \Rightarrow w = \sum_{k=1}^{N}(\alpha_k - \alpha_k^*)\varphi(x_k)$$

$$\frac{\partial \mathcal{L}}{\partial b} = -\sum_{k=1}^{N}\alpha_k + \sum_{k=1}^{N}\alpha_k^* = 0 \Rightarrow \sum_{k=1}^{N}(\alpha_k - \alpha_k^*) = 0$$

$$\frac{\partial \mathcal{L}}{\partial e_k} = c - \alpha_k - \eta_k = 0, \quad \frac{\partial \mathcal{L}}{\partial e_k^*} = c - \alpha_k^* - \eta_k^* = 0, \quad \frac{\partial \mathcal{L}}{\partial \varepsilon} = c\nu - \sum_{k=1}^{N}\alpha_k - \sum_{k=1}^{N}\alpha_k^* = 0$$
$$\Rightarrow c\nu = \sum_{k=1}^{N}(\alpha_k + \alpha_k^*)$$

Then, the dual Lagrangian has the form:

$$D = \frac{1}{2}w^Tw + c\left(\varepsilon\nu + \sum_{k=1}^{N}(e_k + e_k^*)\right) - \sum_{k=1}^{N}\alpha_k\left(\varepsilon + e_k + w^T\varphi(x_k) + b - y_k\right) - \sum_{k=1}^{N}\alpha_k^*\left(\varepsilon + e_k^* - w^T\varphi(x_k) - b + y_k\right)$$
$$- \sum_{k=1}^{N}\left((c - \alpha_k)e_k + (c - \alpha_k^*)e_k^*\right)$$
$$= \frac{1}{2}w^Tw + c\varepsilon\nu - \sum_{k=1}^{n}\alpha_k\left(\varepsilon + w^T\varphi(x_k) - y_k\right) - \sum_{k=1}^{N}\alpha_k^*\left(\varepsilon - w^T\varphi(x_k) + y_k\right)$$
$$= -\frac{1}{2}w^Tw + \varepsilon c\nu - \varepsilon\sum_{k=1}^{N}(\alpha_k + \alpha_k^*) + \sum_{k=1}^{N}y_k(\alpha_k - \alpha_k^*)$$
$$= -\frac{1}{2}\sum_{k,l=1}^{N}(\alpha_k - \alpha_k^*)\varphi(x_k)^T\varphi(x_l)(\alpha_l - \alpha_l^*) + \sum_{k=1}^{N}y_k(\alpha_k - \alpha_k^*)$$

Finally:

$$\min_{\alpha, \alpha_k^*} D = -\frac{1}{2} \sum_{k,l=1}^{N} (\alpha_k - \alpha_k^*) \varphi(x_k)^T \varphi(x_l) (\alpha_l - \alpha_l^*) + \sum_{k=1}^{N} y_k (\alpha_k - \alpha_k^*)$$

$$\text{s.t} \quad \sum_{k=1}^{N} (\alpha_k - \alpha_k^*) = 0$$

$$\sum_{k=1}^{N} (\alpha_k + \alpha_k^*) = C\nu$$

$$\alpha_k, \alpha_k^* \in [0, C]$$

MAPE - ν Support vector regression

$$\min_{w,b,e_k,e_k^*} \ \frac{1}{2} w^T w + C\left(\nu\varepsilon + \sum_{k=1}^{N}(e_k + e_k^*)\right)$$

$$\text{s.t} \quad \frac{y_k - w^T\varphi(x_k) - b}{y_k} \le \varepsilon + e_k \qquad e_k, e_k^* \geq 0$$

$$\frac{w^T\varphi(x_k) + b - y_k}{y_k} \le \varepsilon + e_k^*$$

Lagrangian:

$$L = \frac{1}{2} w^T w + C\left(\nu\varepsilon + \sum_{k=1}^{N}(e_k + e_k)\right) - \sum_{k=1}^{N}\alpha_k\left((\varepsilon + e_k)y_k + w^T\varphi(x_k) + b - y_k\right)$$

$$- \sum_{k=1}^{N}\alpha_k^*\left((\varepsilon + e_k^*)y_k - w^T\varphi(x_k) - b + y_k\right) - \sum_{k=1}^{n}\left(\eta_k e_k + \eta_k^* e_k^*\right)$$

Saddle point:

$$\frac{\partial L}{\partial w} = w - \sum_{k=1}^{N}\alpha_k\varphi(x_k) + \sum_{k=1}^{N}\alpha_k^*\varphi(x_k) = 0 \Rightarrow w = \sum_{k=1}^{N}(\alpha_k - \alpha_k^*)\varphi(x_k)$$

$$\frac{\partial L}{\partial b} = -\sum_{k=1}^{N}\alpha_k + \sum_{k=1}^{N}\alpha_k^* = 0 \Rightarrow \sum_{k=1}^{N}(\alpha_k - \alpha_k^*) = 0$$

$$\frac{\partial L}{\partial e_k} = C - \alpha_k y_k - \eta_k = 0, \quad \frac{\partial L}{\partial e_k^*} = C - \alpha_k^* y_k - \eta_k^* = 0$$

$$\frac{\partial L}{\partial \varepsilon} = C\nu - \sum_{k=1}^{N}\alpha_k y_k - \sum_{k=1}^{N}\alpha_k^* y_k = 0 \Rightarrow C\nu = \sum_{k=1}^{N}(\alpha_k + \alpha_k^*)y_k$$

Thus:

$$D = \frac{1}{2} w^T w + C\left(\nu\varepsilon + \sum_{k=1}^{N}(e_k + e_k^*)\right) - \sum_{k=1}^{N}\alpha_k\left((\varepsilon + e_k)y_k + w^T\varphi(x_k) + b - y_k\right)$$

$$- \sum_{k=1}^{N}\alpha_k^*\left((\varepsilon + e_k^*)y_k - w^T\varphi(x_k) - b + y_k\right) - \sum_{k=1}^{N}\left((C - \alpha_k y_k)e_k + (C - \alpha_k^* y_k)e_k^*\right)$$

$$= -\frac{1}{2} w^T w + \varepsilon C\nu - \sum_{k=1}^{N}\alpha_k(\varepsilon - 1)y_k - \sum_{k=1}^{N}\alpha_k^*(\varepsilon + 1)y_k$$

$$= -\frac{1}{2} w^T w + \varepsilon C\nu - \varepsilon\sum_{k=1}^{N}(\alpha_k + \alpha_k^*)y_k + \sum_{k=1}^{N}(\alpha_k - \alpha_k^*)y_k$$

$$= -\frac{1}{2}\sum_{k,l=1}^{N}(\alpha_k - \alpha_k^*)\varphi(x_k)^T\varphi(x_l)(\alpha_l - \alpha_l^*) + \sum_{k=1}^{N}(\alpha_k - \alpha_k^*)y_k$$

A LS-SVM approach with percentage-like error

$$\min \quad \frac{1}{2}w^T w + \Gamma \frac{1}{2} \sum_{k=1}^{N} \frac{e_k^2}{y_k}$$

Percentage Proxy

$$\text{s.t.} \quad y_k = w^T \varphi(x_k) + b + e_k$$

Lagrangian

$$\mathcal{L}(w, b, \Gamma, \alpha) = \frac{1}{2} w^T w + \Gamma \frac{1}{2} \sum_{k=1}^{N} \frac{e_k^2}{y_k} - \sum_{k=1}^{N} \alpha_k \left(y_k - w^T \varphi(x_k) - b + e_k\right)$$

The conditions for optimality yield:

$$\frac{\partial \mathcal{L}}{\partial w} = 0 \Rightarrow w - \sum_{k=1}^{N} \alpha_k \varphi(x_k) = 0 \Rightarrow w = \sum_{k=1}^{N} \alpha_k \varphi(x_k)$$

$$\frac{\partial \mathcal{L}}{\partial e_k} = 0 \Rightarrow \Gamma \frac{e_k}{y_k} - \alpha_k = 0 \Rightarrow e_k = \frac{\alpha_k y_k}{\Gamma}$$

$$\frac{\partial \mathcal{L}}{\partial b} = 0 \Rightarrow \sum_{k=1}^{N} \alpha_k = 0$$

$$\frac{\partial \mathcal{L}}{\partial \alpha_k} = 0 \Rightarrow w^T \varphi(x_k) + b + e_k - y_k = 0$$

How to solve this system of equations? ? ? ? ? ? ? ?

That's the main result

What if: $\sum_{l=1}^{N} \alpha_l \varphi(x_l)^T \varphi(x_k) + b + \frac{\alpha_k y_k^2}{\Gamma} - y_k = 0$

$$= \left(\alpha_1 \varphi(x_1)^T + \alpha_2 \varphi(x_2)^T + \dots + \alpha_N \varphi(x_N)^T\right) \varphi(x_k) + b + \frac{\alpha_k y_k^2}{\Gamma} - y_k = 0$$

$$= \alpha_1 \varphi(x_1)^T \varphi(x_k) + \dots + \alpha_N \varphi(x_N)^T \varphi(x_k) + b + \frac{\alpha_k y_k^2}{\Gamma} - y_k = 0$$

Thus, for $k = 1, \dots, N$

$$\left.\begin{array}{l}
\alpha_1 \varphi(x_1)^T \varphi(x_1) + \dots + \alpha_N \varphi(x_N)^T \varphi(x_1) + b + \frac{\alpha_1 y_1^2}{\Gamma} - y_1 = 0 \\
\alpha_1 \varphi(x_1)^T \varphi(x_2) + \dots + \alpha_N \varphi(x_N)^T \varphi(x_2) + b + \frac{\alpha_2 y_2^2}{\Gamma} - y_2 = 0 \\
\vdots \\
\alpha_1 \varphi(x_1)^T \varphi(x_N) + \dots + \alpha_N \varphi(x_N)^T \varphi(x_N) + b + \frac{\alpha_N y_N^2}{\Gamma} - y_N = 0
\end{array}\right\} \text{KKT system}$$

Let: $\Omega_{lk} = [K(x_l, x_k)]$, $1_v = [1, \dots, 1]^T$, $Y_\Gamma = \mathrm{diag}(y_1^2/\Gamma, y_2^2/\Gamma, \dots, y_N^2/\Gamma)$

$\alpha = [\alpha_1, \alpha_2, \dots, \alpha_N]^T$, $y = [y_1, y_2, \dots, y_N]^T$

Then, the KKT system has the form:

$$(\Omega + Y_\Gamma)\alpha + b 1_v = y$$

On the other hand:

$$\sum_{k=1}^{N} \alpha_k = 0 \Rightarrow 1_v^T \alpha = 0$$

Joining both equations:

$$\left[\begin{array}{c|c} 0 & 1_v^T \\ \hline 1_v & \Omega + Y_\Gamma \end{array}\right] \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}$$

Vector formulation of the LS-SVM

$$\min_{w,b,e} \ \frac{1}{2} w^T w + \Gamma \frac{1}{2} (e \circ z)^T (e \circ z)$$

$$\text{s.t.} \quad y = \Phi w + b 1_v + e$$

Lagrangian:

$$L = \frac{1}{2} w^T w + \Gamma \frac{1}{2} (e \circ z)^T (e \circ z) - \alpha^T (y - \Phi w - b 1_v - e)$$

$$\frac{\partial L}{\partial w} = w - \alpha^T \Phi = 0 \Rightarrow w = \alpha^T \Phi$$

$$\frac{\partial L}{\partial b} = \alpha^T 1_v = 0$$

$$\frac{\partial L}{\partial e} = \Gamma (e \circ z) \circ z - \alpha = 0 \Rightarrow e = \frac{1}{\Gamma} \alpha \circ (Y \circ Y)$$

$$\frac{\partial L}{\partial \alpha} = y - \phi w - b 1_v - e = 0$$