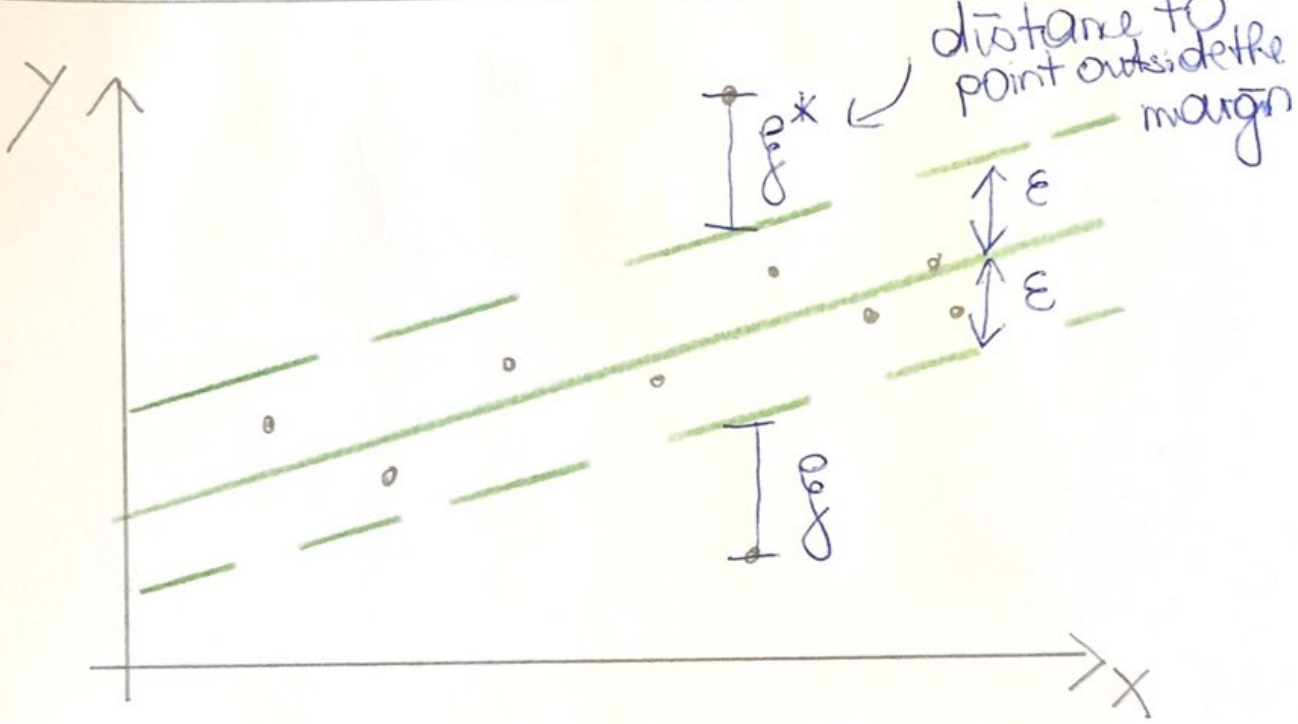
Understanding the slack variables in SVR



Optimization: find narrowest tube while minimizing the ~~resid~~ forecast error

find narrowest tube: $\min\left[\frac{1}{2}\|\omega\|^2\right]$

forecast error: for instance MSE

$$\begin{cases} 0 & \text{if } |y-\hat{y}| \leq \varepsilon \\ (|y-\hat{y}|-\varepsilon)^2 & \text{else} \end{cases}$$

Soft-margin: ξ and ξ^* help to tolerate points outside the margin

$$\min \left[\frac{1}{2} \|w\|^2 \right] + C \sum_{i=1}^{N} \left[\xi_i + \xi_i^* \right]$$

regularization, larger $C \rightarrow$ more weight to minimizing the error

Inequality constraints:

$$y_i - w^T x_i \leq \varepsilon + \xi_i^* \qquad i = 1 \ldots N$$

$$y_i - w^T x_i \geq \varepsilon + \xi_i \qquad i = 1 \ldots N$$

Next step: Rewriting optimization problem with Lagrangian and then deriving the dual form.