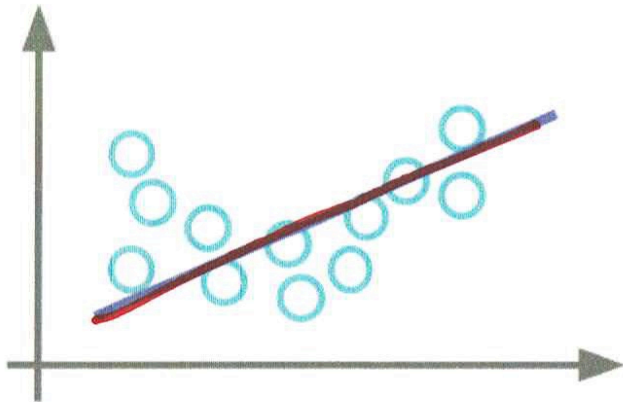
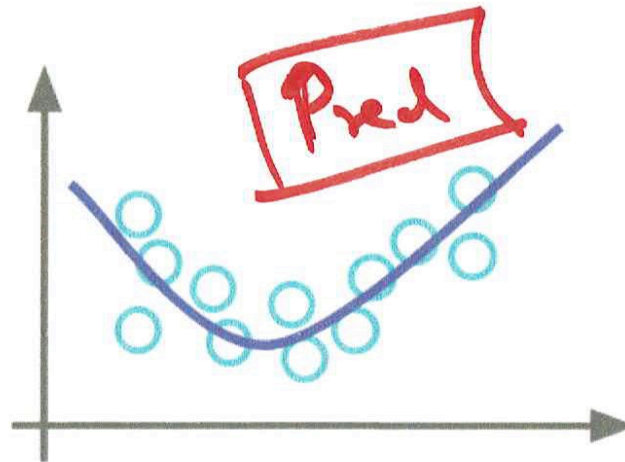


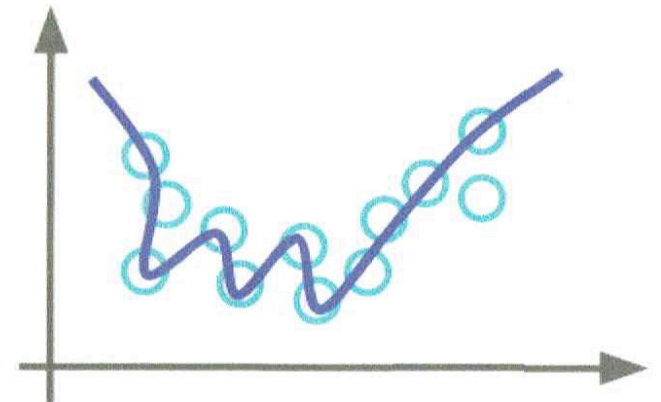
$$\text{Data} = \underbrace{\text{Inf}}_{\substack{\uparrow \\ \text{model}}} + \text{Noise} \quad \times^T$$



Under



✓



Over

L_1 or L_2 Reg

$L_1 \rightarrow$ Lasso ←

$L_2 \rightarrow$ Ridge

min \downarrow

\uparrow

$L(y, \hat{y}) + \lambda(\text{penalty})$

$\rightarrow \underline{\sum |w_i|}$ Lasso

$\rightarrow \underline{\sum (w_i)^2}$ Ridge

$$\omega^j = \omega^{j-1} - 2 \underline{\nabla_{\omega} l(\omega)}$$

SGD
with momentum

$$\underline{\omega^j = \omega^{j-1} - 2 g^j}$$

$$g^j = \nabla_{\omega} l(\omega)$$

add
momentum

$$\omega^j = \omega^{j-1} - 2 \underline{(\alpha g^j + (1-\alpha) g^{j-1})}$$