L2 loss

$$L_2 = |f(x) - Y|^2$$
$$L_2' = 2f'(x)(f(x) - Y) \quad \text{..... (1)}$$

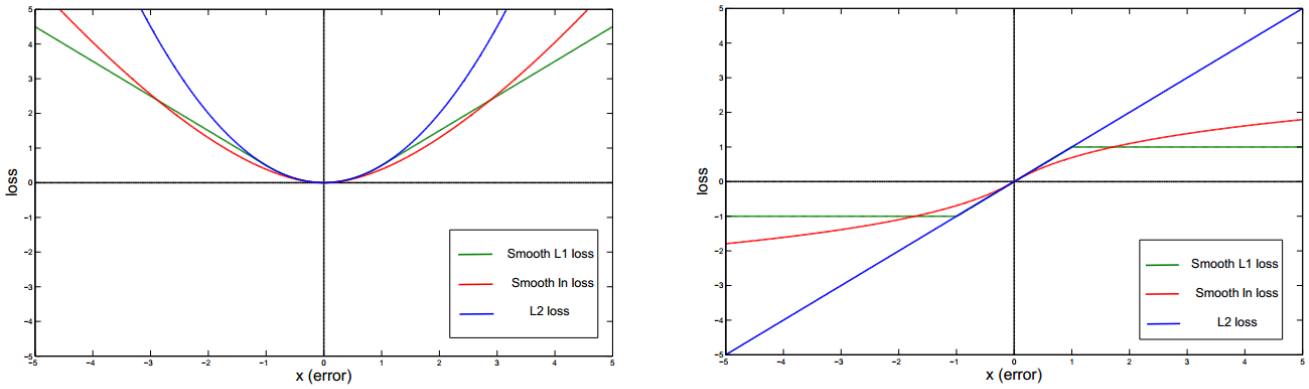
L1 loss

$$L_1 = |f(x) - Y|$$
$$L_1' = \pm f'(x) \quad \text{..... (2)}$$

Smooth L1 loss

$$Smooth \quad L_1 = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & x < -1 \, or \, x > 1 \end{cases}$$

$$Smooth \quad L_1' = \begin{cases} x, & |x| < 1 \\ -1, & x < -1 \\ 1, & x > 1 \end{cases} \quad \text{..... (3)}$$



(a) forward loss functions. (b) backward deviation functions.

L1loss在零点导数不连续，可能影响收敛。用smooth L1进行改进，而且smooth L1对离群点不敏感，梯度变化相对更小，训练时不容易跑飞

L2的缺点是对离群点(outliers)敏感，抗噪声等干扰比较弱。