$$X=\begin{pmatrix}X_1^T\\ \vdots\\ X_n^T\end{pmatrix}=(X_{ij})_{n\times p},\quad \beta=\begin{pmatrix}\beta_1\\ \vdots\\ \beta_p\end{pmatrix}\quad Y=\begin{pmatrix}Y_1\\ \vdots\\ Y_p\end{pmatrix}.$$

$$L(\beta)=\frac{1}{2}\|Y-X\cdot\beta\|_2^2=\frac{1}{2}\sum_{i=1}^{n}(Y_i-X_i^T\beta)^2=\frac{1}{2}\sum_{i=1}^{n}\left(Y_i-\sum_{j=1}^{p}X_{ij}\cdot\beta_j\right)^2.$$

$$\frac{\partial L(\beta)}{\partial\beta_k}=\sum_{i=1}^{n}\left(Y_i-\sum_{j=1}^{p}X_{ij}\cdot\beta_j\right)(-X_{ik})=\sum_{i=1}^{n}\left(\sum_{j=1}^{p}X_{ij}\beta_j-Y_i\right)X_{ik}.$$

$$\begin{pmatrix}\frac{\partial L(\beta)}{\partial\beta_1}\\ \vdots\\ \frac{\partial L(\beta)}{\partial\beta_p}\end{pmatrix}=\sum_{i=1}^{n}\left(\sum_{j=1}^{p}X_{ij}\beta_j-Y_i\right)\cdot\vec{X}_i=0$$

$$(Y^T\cdot X\cdot\beta)^T$$

$$L(\beta)=\frac{1}{2}\cdot(Y-X\cdot\beta)^T\cdot(Y-X\beta)=\frac{1}{2}\cdot(Y^TY-Y^T\cdot X\cdot\beta-\beta^T\cdot X^T\cdot Y+\beta^T\cdot X^T\cdot X\cdot\beta)$$

$$=\frac{1}{2}\cdot Y^T\cdot Y-\beta^T\cdot(X^T\cdot Y)+\frac{1}{2}\cdot\beta^T\cdot X^T\cdot X\cdot\beta.$$

$$\boxed{\frac{\partial(\beta^T\cdot a)}{\partial\beta}=\begin{pmatrix}\frac{\partial(\beta^T\cdot a)}{\partial_i[\beta_i]}\end{pmatrix}_{p\times 1}=(a_i)_{p\times 1}=a.}$$

$a^T\beta$

$$\frac{\partial L(\beta)}{\partial\beta}=\underset{p\times p}{(X^T\cdot X)}\cdot\underset{p\times 1}{\beta}-\underset{p\times 1}{X^T\cdot Y}$$

$$\hat{\beta}=(X^T\cdot X)^{-1}\cdot(X^T\cdot Y)\quad\left(\frac{\partial^2 L(\beta)}{\partial\beta\cdot\partial\beta^T}=X^T\cdot X>>0\right)$$

$$\frac{\partial\left[\frac{1}{2}\beta^T A\beta\right]}{\partial\beta}=\frac{1}{2}A\cdot\beta+\frac{1}{2}\cdot(\beta^T A)^T=\frac{1}{2}A\cdot\beta+\frac{1}{2}\cdot A^T\cdot\beta=\left(\frac{1}{2}A+\frac{1}{2}A^T\right)\cdot\beta.$$

$$\beta^T\cdot A\cdot\beta=\beta^T\cdot\left(\frac{A+A^T}{2}\right)\cdot\beta$$

$\varepsilon_1,\dots,\varepsilon_n$ iid $\sim U[-a,a]$

$Y_i=X_i^T\beta+\varepsilon_i$

$|X_i:\quad Y_i\sim[X_i^T\beta-a,\ X_i^T\beta+a]$

$$L(\beta,a)=\prod_{i=1}^{n}\frac{1}{2a}\cdot I(Y_i-X_i^T\beta\in(-a,a))$$

$a\downarrow$

MLE:

$\boxed{(X, Y) \sim F(x, y)}$ $\quad X \to g(X) \approx Y$

$E(Y - g(X))^2$

$= E(Y - f(X))^2 + E(f(X) - g(X))^2$,

$F(x, y) \Longrightarrow f(x) = E(Y \mid X = x)$ 回归函数、条件期望、

$X \approx x$

$f(x) = E(Y \mid X = x) \Longleftarrow KNN$

$$\hat{f}(x) = \frac{\sum_{i=1}^{n} Y_i \cdot I(\|X_i - x\|_2^2 \le \varepsilon)}{\sum_{i=1}^{n} I(\|X_i - x\|_2^2 \le \varepsilon)}$$