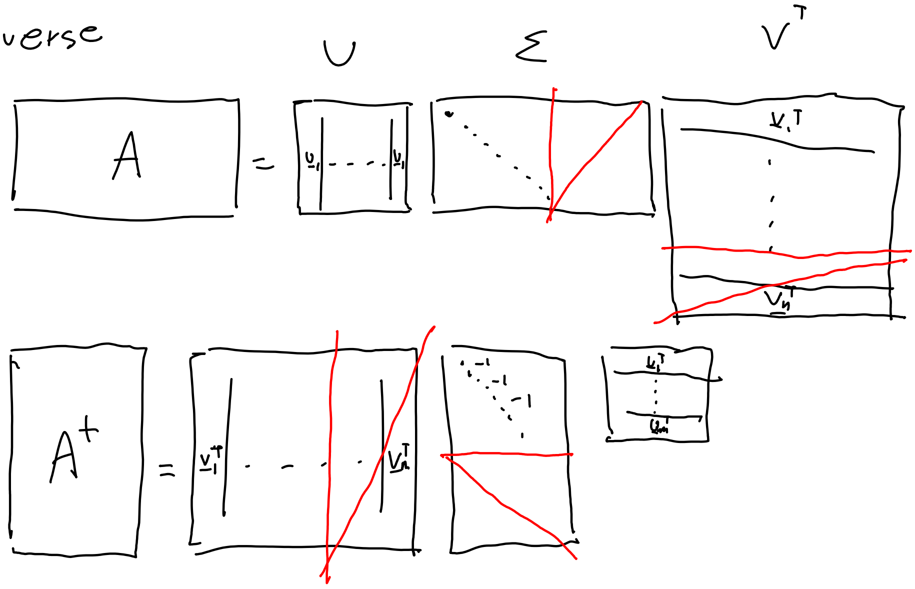
Moore Penrose Pseudo-inverse

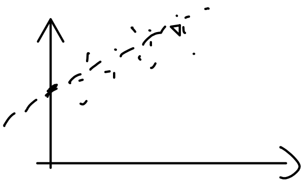
$$U \Sigma V^T = A$$

$$\sigma_j^+ = \begin{cases} 1/\sigma_j & , \sigma_j > 0 \\ 0 & , \sigma_j = 0 \end{cases}$$

$$A^+ = V \Sigma^+ U^T$$



Least Squares R.



$(x_i, y_i) \quad i = 1, \dots, n \qquad y_i = m x_i + q$

$$(\hat{m}, \hat{q}) := \underset{m,q}{\operatorname{argmin}} \sum_{i=1}^{n} \left(y_i - (m x_i + q)\right)^2$$

$$\begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \begin{bmatrix} m \\ q \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \qquad \Phi = \begin{bmatrix} \varphi(x_1) \\ \vdots \\ \varphi(x_n) \end{bmatrix}$$

φ Feature map $\quad \varphi(x) = \begin{pmatrix} x \\ 1 \end{pmatrix}$, $\varphi(x) = x$

$\Phi \quad \underline{w} \quad y \rightsquigarrow y = \Phi \underline{w} \qquad \underline{w}_{LS} := \underset{\underline{w}}{\operatorname{argmin}} \|y - \Phi \underline{w}\|^2$

$$\underline{w}_{LS} = (\Phi^T \Phi)^{-1} \Phi^T y = \Phi^+ y = V \Sigma^+ U^T y \qquad \Phi = U \Sigma V^T$$

Ridge Regression

$$\underline{w}_{RR} := \underset{\underline{w}}{\operatorname{argmin}} \|y - \Phi \underline{w}\|_2^2 + \overbrace{\lambda \|\underline{w}\|_2^2}^{\text{regularization}}$$

$$\underline{w}_{RR} = \Phi^T \underbrace{(\Phi \Phi^T + \lambda I)^{-1} y}_{\alpha}$$

$$\begin{cases} y^{new} = \varphi(\underline{x}^{new}) \cdot \underline{w}_{RR} = \varphi(\underline{x}^{new}) \Phi^T \underline{\alpha} = \sum_{i=1}^{n} \alpha_i \varphi(\underline{x}^{new}) \cdot \varphi(\underline{x}_i) \\ \underline{w}_{RR} = \Phi^T \underline{\alpha} \\ \underline{\alpha} = (\Phi \Phi^T + \lambda I)^{-1} y \end{cases}$$

$$= \sum_{i=1}^{n} \alpha_i K(\underline{x}^{new}, \underline{x}_i)$$

$$(\Phi \Phi^T)_{ij} = \underbrace{\varphi(\underline{x}_i) \cdot \varphi(\underline{x}_j)}_{=: K(\underline{x}_i, \underline{x}_j)} =: K_{ij}$$

Kernel Trick

Kernel Regression

$$y^{new} = \sum_{i=1}^{n} \alpha_i K(\underline{x}^{new}, \underline{x}_i)$$

$$\underline{\alpha} = (K + \lambda I)^{-1} y$$

$$\varphi(x) = \begin{pmatrix} x \\ 1 \end{pmatrix}$$

$$K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j) = \begin{pmatrix} x_i \\ 1 \end{pmatrix} \cdot \begin{pmatrix} x_j \\ 1 \end{pmatrix} = x_i x_j + 1$$