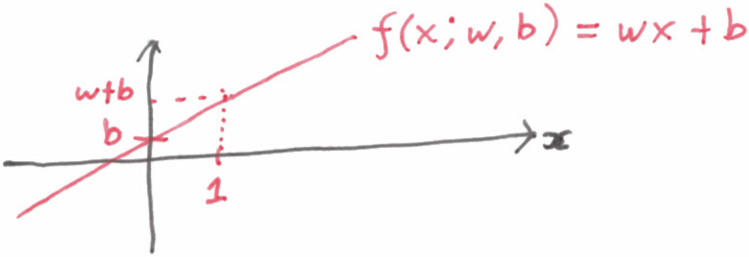
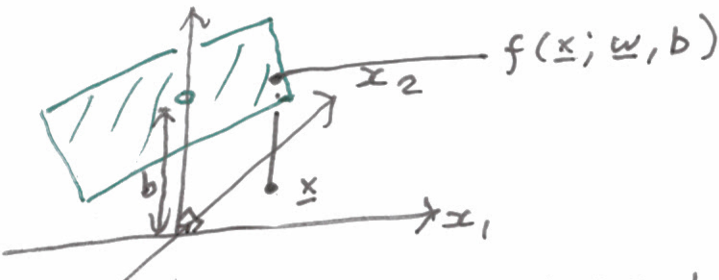
Linear Functions

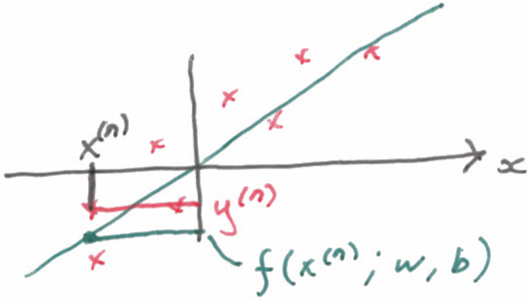




$$f(\underline{x};\underline{w},b) = w_1x_1 + w_2x_2 + b$$
$$= \underline{w}^T\underline{x} + b$$

Data

$$\{(\underline{x}^{(n)}, y^{(n)})\}_{n=1}^{N}$$



Residual $y^{(n)} - f(x^{(n)}; w, b)$

$$\underline{y} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(N)} \end{bmatrix} \qquad X = \begin{bmatrix} - & \underline{x}^{(1)T} & - \\ - & \underline{x}^{(2)T} & - \\ & \vdots & \\ x_1^{(N)} & x_2^{(N)} \cdots & x_D^{(N)} \end{bmatrix}$$

$N \times 1$ "matrix" $\qquad N \times D$ matrix $\qquad$ D-dimensional regression

Python vector y is $(N,)$

Numpy: y[:, None] is $(N,1)$ array

$$\underline{f} = \begin{bmatrix} f(\underline{x}^{(1)}; \underline{w}, b) \\ \vdots \\ f(\underline{x}^{(N)}; \underline{w}, b) \end{bmatrix}$$

Least squares fitting

Minimize $\displaystyle\sum_{n=1}^{N} \left(y^{(n)} - f(\underline{x}^{(n)}; \underline{w}, b) \right)^2$

Minimize: $(\underline{y} - \underline{f})^T (\underline{y} - \underline{f})$

Models with zero intercept $(b=0)$

$f(\underline{x}; \underline{w}) = \underline{w}^T \underline{x} = \underline{x}^T \underline{w}$

$\underset{N\times 1}{\underline{f}} = \underset{N\times D}{X} \; \underset{D\times 1}{\underline{w}} \approx \underline{y}$

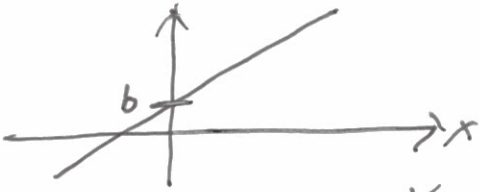
"Linear map"

$g(\underline{x} + \underline{z}) = g(\underline{x}) + g(\underline{z})$

$g(c\underline{x}) = c\, g(\underline{x})$

Matlab: `w_fit = X \ y;`

Python `= np.linalg.lstsq(X, y)[0]`



$$X' = \begin{bmatrix} 1 & \underbrace{- \; \underline{x}^{(1)T} \; -}^{X} \\ 1 & \vdots \\ \vdots & \vdots \\ 1 & - \; \underline{x}^{(N)T} \; - \end{bmatrix} \qquad N \times (D+1)$$

$$\underline{w}' = \operatorname{argmin} \| \underline{y} - X'\underline{w}' \|^2$$

Fit $\underline{y}$ with $\underline{f} = X'\underline{w}' = w_1' + X\underline{w}'_{2:D+1}$

$$= b + X\underline{w}$$

or a $N \times 1$ vector with every element $= w_1' = b$

$$\underline{f} = \underline{\Phi}\,\underline{w}$$

$\underline{\Phi}$: $N \times K$ any representation of data.

Each row is a datapoint

$$\underline{\Phi} = \begin{bmatrix} \text{---} & \underline{\phi}(\underline{x}^{(1)})^T & \text{---} \\ & \vdots & \\ \text{---} & \underline{\phi}(\underline{x}^{(N)})^T & \text{---} \end{bmatrix}$$

Example

$$\underline{\phi}(x) = [1 \;\; x \;\; x^2 \;\; x^3]^T$$

Could have $x_1 x_2^3 x_3$ in 3D

Fit $\underline{y} \approx \underline{f} = \underline{\Phi}\,\underline{w}$

$$f(x) = w_1 + w_2 x + w_3 x^2 + w_4 x^3$$