General Regression models

Dr. Pritam Anand.
Assistant Professor,
DA-IICT, Gandhinagar.

Multiple Regression model working with k variables

x_1	x_2	x3	--	x_k	y Balance (thousand dollar)
32	0.550798	283	--	2	5.651202
22	0.708148	483	-	3	7.321263
45	0.290905	514	-	4	5.167304
78	0.510828	681	-	3	5.609367
54	0.892947	357	-	2	9.406379
39	0.896293	569	-	4	9.379439
42	0.125585	259	-	2	2.734997
51	0.207243	512	-	2	4.876649
21	0.051467	266	-	5	3.584138
19	0.44081	491	-	3	5.437239

Multiple Regression model working with k variables

Linear Function:-

$f(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \ldots + \beta_k x_k$

x_1 x_2 x_k

For given Training Set $T = \{ (x_{11}, x_{12, \ldots}, x_{1k}, y_1), (x_{21}, x_{22, \ldots}, x_{2k}, y_2), \ldots, (x_{n1}, x_{n2, \ldots}, xnk, y_n)\}$, we solve

Min $J(\beta_k, \ldots, \beta_4, \beta_3, \beta_2, \beta_1, \beta_0) = \sum_{i=1}^{n} (yi - (\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \ldots + \beta_k x_k))^2$

$$u = \begin{bmatrix} \beta_k \\ \\ \\ \\ \beta_2 \\ \beta_1 \\ \beta_0 \end{bmatrix} \qquad A = \begin{bmatrix} x_{11} & x_{21} & -- & x_{1k} & 1 \\ x_{21} & x_{22} & -- & x_{2k} & 1 \\ \\ \\ x_{n1} & x_{n2} & -- & x_{nk} & 1 \end{bmatrix} \qquad Y = \begin{bmatrix} y_1 \\ y_2 \\ \\ \\ y_n \end{bmatrix}$$

The Least Square problem reduces to

$$Min_{(u)} \quad J(\mathrm{u}) = \quad (Y - Au)^T (Y - Au)$$

$$u = \begin{bmatrix} \beta_k \\ \\ \\ \\ \beta_2 \\ \beta_1 \\ \beta_0 \end{bmatrix} = (A^T A)^{-1} \; A^T Y$$

Change of Notation

- We are working with n number of variables. So each independent variable *X* is in R^n

- Our training set $T = \{(\boldsymbol{x}_i, yi): \boldsymbol{x}_i \in Rn \; and \; yi \in R, i = 1,2, \dots l\}$.

  n

  i i

- The linear function in R^n is given by

$$f(\boldsymbol{x}) = w_1x_1 + w_2x_2 + .. + wnxn + b = wTx + b.$$

T

$$W = \begin{bmatrix} w_1 \\ w_2 \\ - \\ w_n \end{bmatrix}$$

Linear Regression model

$$f(\boldsymbol{x}) = w_1x_1 + w_2x_2 + .. + wnxn + b = wTx + b.$$

- For simple Least Square Regression model we need to minimize

$$\min_{(w \in Rn,\, b \in R)} \sum_{i=1}^{l} (yi - (w^T\boldsymbol{x}_i + b))$$

$$U = \begin{bmatrix} w_n \\ \\ \\ \\ w_2 \\ w_1 \\ b \end{bmatrix} \qquad A = \begin{bmatrix} x_{11} & x_{21} & -- & x_{1k} & 1 \\ x_{21} & x_{22} & -- & x_{2k} & 1 \\ \\ \\ x_{n1} & x_{n2} & -- & x_{nk} & 1 \end{bmatrix} \qquad Y = \begin{bmatrix} y_1 \\ y_2 \\ \\ \\ y_n \end{bmatrix}$$

The Least Square problem reduces to

$$Min_{(u)} \quad J\ \mathrm{u}) = \quad (Y - Au)^T (Y - Au)$$

$$\boxed{\mathrm{u}== \begin{bmatrix} w_n \\ \\ \\ w_2 \\ w_1 \\ b \end{bmatrix} = (A^T A)^{-1} A^T \mathrm{Y}}$$

Regularized Linear Regression model

$$f(\boldsymbol{x}) = w_1x_1 + w_2x_2 + .. + wnxn + b = wTx + b.$$

- For simple Least Square Regression model we need to minimize

$$\min_{(w \in Rn,\, b \in R)} \frac{\lambda}{2} \mathrm{w}^{\mathrm{T}}\mathrm{w} + \sum_{i=1}^{l} (yi - (w^T \boldsymbol{x}_i + b))$$

$$U = \begin{bmatrix} w_n \\ \\ \\ \\ w_2 \\ w_1 \\ b \end{bmatrix} \qquad A = \begin{bmatrix} x_{11} & x_{21} & -- & x_{1k} & 1 \\ x_{21} & x_{22} & -- & x_{2k} & 1 \\ \\ \\ x_{n1} & x_{n2} & -- & x_{nk} & 1 \end{bmatrix} \qquad Y = \begin{bmatrix} y_1 \\ y_2 \\ \\ \\ y_n \end{bmatrix}$$

The Least Square problem reduces to

$$Min_{(u)} \quad J\,u) = \quad (Y - Au)^T\,(Y - Au) + \frac{\lambda}{2} u^T u$$

$$\boxed{u = \begin{bmatrix} w_n \\ \\ \\ w_2 \\ w_1 \\ b \end{bmatrix} = (A^T A + \lambda I)^{-1} A^T Y}$$

A quadratic function in R^n

$$\frac{(n+2)!}{n!\,2!} = \frac{(n+1)(n+2)}{2}$$

$$w_m x_n^2 \rightarrow \varphi_m(\mathbf{x})$$

$$+ w_{m-1} x_{n-1}^2 \rightarrow \varphi_{m-1}(\mathbf{x})$$

$$-$$

$$+ w_j x_1^2$$

$$+ w_k x_n x_{n-1}$$

$$+ x_n x_{n-2}$$

$$=$$

$$-$$

$$+ x_n$$

$$+ x_{n-1}$$

$$=$$

$$-$$

$$+ w_1 x_1 - \varphi_1(x)$$

$$b\,1 - \varphi_0(x)$$

$$= w_m \varphi_m(\mathbf{x}) + w_{m-1}\varphi_{m-1}(\mathbf{x}) + \ldots + w_1\varphi_1(\mathbf{x}) + b\varphi_0(\mathbf{x})$$

$$f(x) = w^T \varphi(x) + b$$

$$\mathbf{w} = \begin{bmatrix} w_m \\ w_{m-1} \\ \vdots \\ w_1 \end{bmatrix} \qquad \varphi(x) = \begin{bmatrix} \varphi_m(x) \\ \varphi_{m-1}(x) \\ \vdots \\ \varphi_1(x) \\ 1 \end{bmatrix}$$

$$= \boxed{\sum_{i=1}^{m} w_i \varphi_i(x) + b}$$

$$\mathbf{w} = \begin{bmatrix} w_m \\ w_{m-1} \\ \vdots \\ w_1 \\ b \end{bmatrix} \qquad w^T \varphi(x)$$

$$f(x) = w^T \phi(x)$$

$$w = \begin{bmatrix} w_m \\ w_{m-1} \\ \vdots \\ w_1 \\ b \end{bmatrix} \qquad \phi(x) = \begin{bmatrix} \phi_m(x) \\ \phi_{m-1}(x) \\ \vdots \\ \phi_1(x) \\ 1 \end{bmatrix}$$

$$\rightarrow \; w^T x + b$$

$$w^T \phi(x) + b$$

$x_1^3 \Rightarrow \phi_m(x)$ —

$x_2^3 - \phi_{m-1}(x)$ —

—

$x_n^3 - \phi_{m-n}(x)$ —

$x_n^2 x_1 - \phi_{m-n-1}(x)$ —

$x_n^2 x_2$ —

—

$=$

$$\sum_{j=1}^{m} w_j \phi_j(x) + b$$

$$= w^T \phi(x) + b$$

A quadratic function in R^n

A quadratic function in R^n

$$w_9 x_3^2 + w_8 x_2^2 + w_7 x_1^2 + w_6 x_3 x_1 + w_5 x_3 x_2 + \dots + \dots b$$

$$w_9 \phi_9(x) \quad w_8 \phi_8(x) \quad w_7 \phi_7(x) \quad + \quad b$$

$$= \sum_{j=1}^{9} w_j \phi_j(x) + b = \underline{w^T \phi(x) + b}$$

$$\frac{(m+n)!}{m! \times n!} = \frac{1003}{1000! \times 3!}$$

Polynomial Regression in R^n

$$f(\boldsymbol{x}) = wM\phi_{M(\boldsymbol{x})} + wM\,_{-\,1}\phi_{M\,-\,1(\boldsymbol{x})} + .. + w_1\phi_{1(\boldsymbol{x})} + b = wT\phi(\boldsymbol{x}) + b.$$

$$\phi(\boldsymbol{x}) = \begin{bmatrix} \phi_{M(\boldsymbol{x})} \\ \phi_{M(\boldsymbol{x})} \\ \\ \\ \phi_{1(\boldsymbol{x})} \end{bmatrix}$$

- For regularized Least Square Regression model we need to minimize

$$\min_{(w \in Rn,\, b \in R)} \frac{\lambda}{2} \mathrm{w}^{\mathrm{T}}\mathrm{w} + \sum_{i=1}^{l} (yi - (w^T\phi(\boldsymbol{x}) + b))^2$$

$$\phi(x_1) = \begin{bmatrix} \phi_M(x_1) \\ \phi_{M-1}(x_1) \\ \phi_1(x_1) \end{bmatrix}$$

$$(y_1 - (w^T\phi(x_1)+b))^2 + (y_2 - (w^T\phi(x_2)+b))^2 + \cdots + (y_l - (w^T\phi(x_l)+b))^2$$

$$U = \begin{bmatrix} w_M \\ \\ \\ w_2 \\ w_1 \\ b \end{bmatrix} \qquad A = \begin{bmatrix} \phi_M(\boldsymbol{x}_1) & \phi_{M-1}(\boldsymbol{x}_1) & -- & \phi_1(\boldsymbol{x}_1) & 1 \\ \phi_M(\boldsymbol{x}_2) & \phi_{M-1}(\boldsymbol{x}_2) & -- & \phi_1(\boldsymbol{x}_2) & 1 \\ \\ \\ \phi_M(\boldsymbol{x}_l) & \phi_{M-1}(\boldsymbol{x}_l) & -- & \phi_1(\boldsymbol{x}_l) & 1 \end{bmatrix} \qquad Y = \begin{bmatrix} y_1 \\ y_2 \\ \\ \\ y_n \end{bmatrix}$$

The Least Square problem reduces to

$$Min_{(u)} \quad J(\mathrm{u}) = (Y - Au)^T (Y - Au) + \frac{\lambda}{2}\mathrm{u}^T\mathrm{u}$$

$$\boxed{\mathrm{u} = \begin{bmatrix} w_M \\ \\ \\ w_2 \\ w_1 \\ b \end{bmatrix} = (A^TA + \lambda I)^{-1} A^T\mathrm{Y}}$$

$$y_1 - \left(\sum_{j=1}^{m} w_j \varphi_j(x_1) + b \right)$$

$$\left(y_1 - \left(w^T \varphi(x_1) + b \right) \right)^2$$

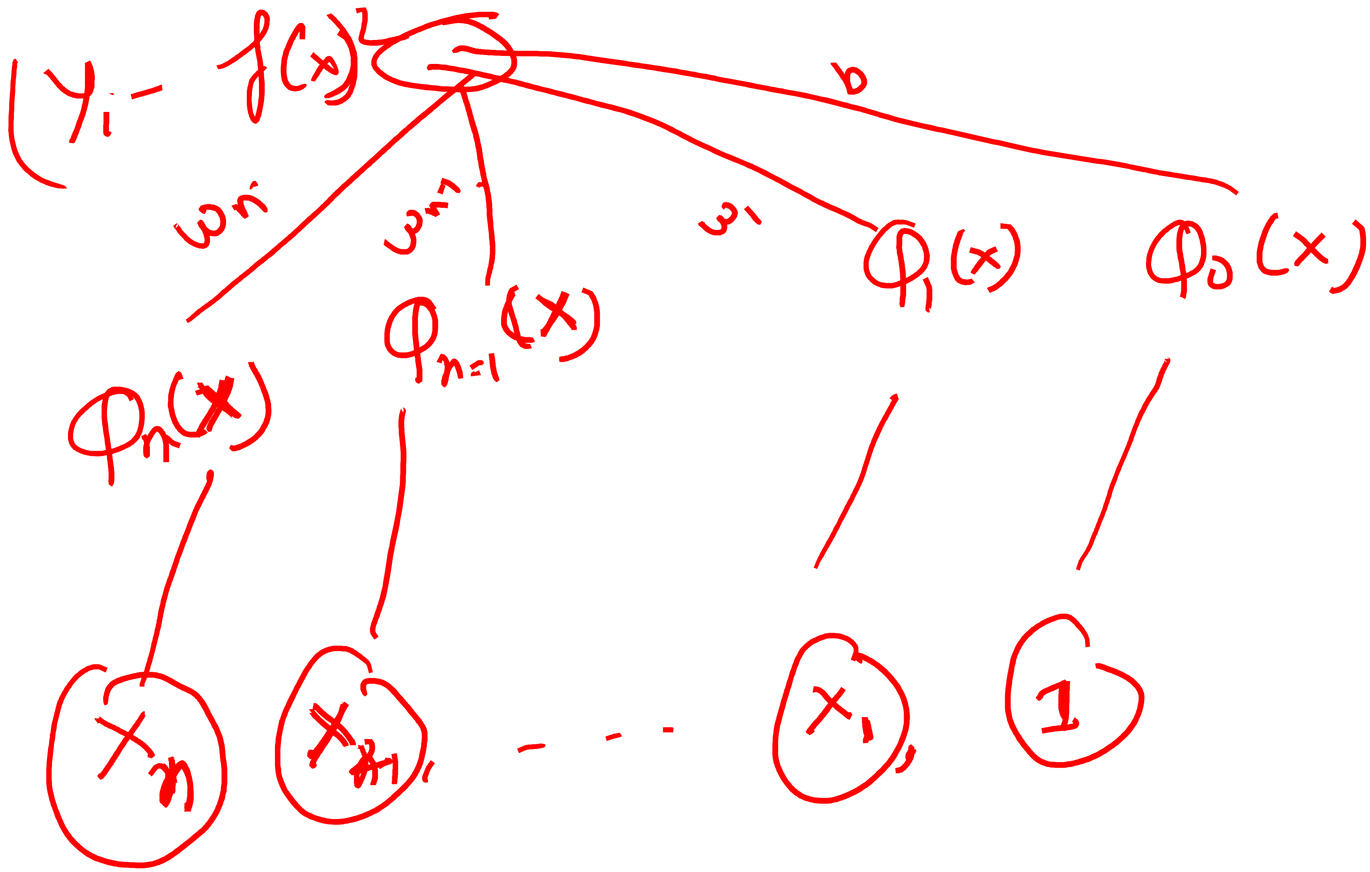
$$+ \left(y_2 - \left(w^T \varphi(x_2) + b \right) \right)^2$$

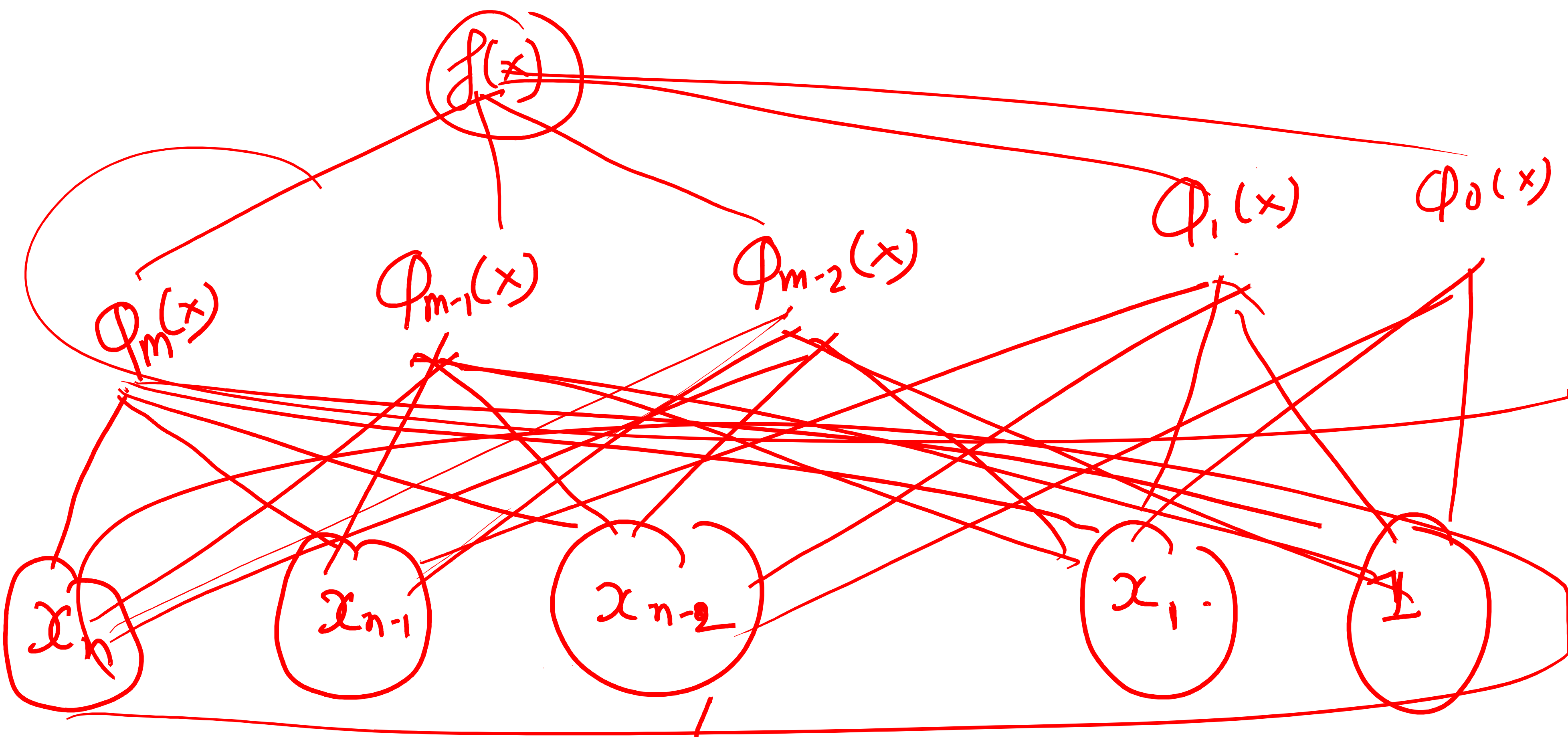
$$+ \left(y_2 - \left(w^T \varphi(x_2) + b \right) \right)^2$$

$$A$$

$$\begin{bmatrix} \phi_m(x_1) & \phi_{m-1}(x_1) & \cdots & \phi_1(x), & 1 \\ \\ \\ \phi_m(x_2) & \phi_{m-1}(x_2) & & \phi_1(x_2), & 1 \end{bmatrix} \overset{Y}{\begin{bmatrix} w_m \\ - \\ - \\ - \\ - \\ w_1 \\ b \end{bmatrix}}$$

$$\begin{bmatrix} w^T\phi(x_1) + b \\ w^T\phi(x_2) + b \\ \cdots \end{bmatrix} \qquad Y - AY$$


(yi - f(x))2
b
wn
wn-1
w1
φn(x)
φn-1(x)
φ1(x)
φ0(x)
xn
xn-1,
x1,
1



$x_n x_{n-2}$

Basis Functions

- The challenge is to find problem specific basis functions which are able to effectively model the true relationship between data

- If we include too few basis functions or unsuitable basis functions, we might not be able to model the true dependency.

- If we include too many basis functions, we need many data points to fit all the unknown parameters

Radial Basis Functions

- Another class are radial basis functions (RBF). Typical representatives are Gaussian basis functions

μ_j

$$\phi_j(x) = \exp\left(-\frac{1}{2sj^2}\,||x - \mathrm{cj}||2\right)$$

σ_j

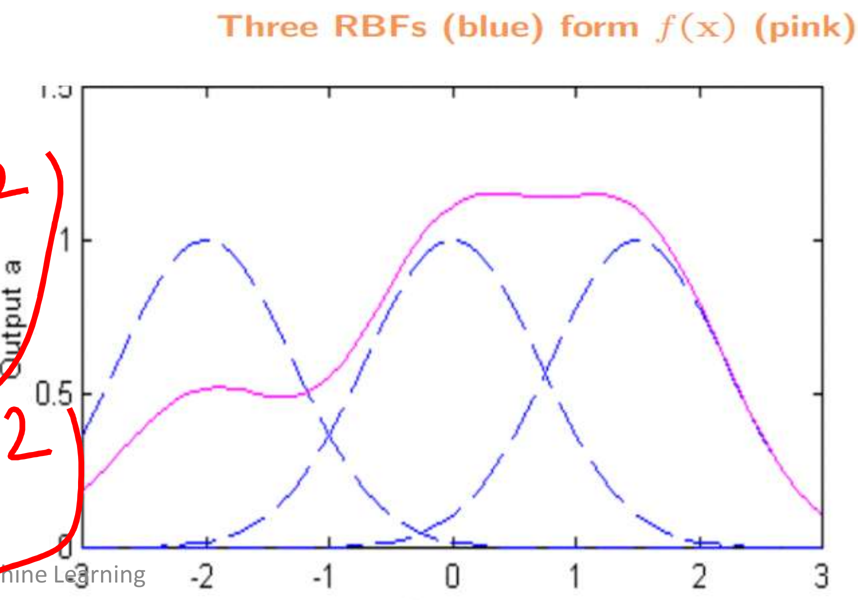
Radial Basis Functions

- Another class are radial basis functions (RBF). Typical representatives are Gaussian basis functions $\phi_j(x) = \exp\ (-\frac{1}{2sj^2}\ ||x - \mathrm{cj}||2)$

$$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$w_3 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x+2}{\sigma}\right)^2} + w_2\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2}\right)$$

$$+\ w_3\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(x-2\right)^2}\right)$$



$$R_1(x) = e^{-\frac{\|x - u_1\|^2}{\sigma^2}}$$

$$x \in R^n$$

$$u_1 \in R^n$$

$$R_k(x) = e^{-\frac{\|x - u_k\|^2}{\sigma^2}}$$

$$\|x - u_k\|^2 = (x - u_k)^T (x - u_k)$$

$$\sigma(x, w) = \frac{1}{1+e^{-(wx+b)}}$$

$$\sigma(x, w) = \sigma((x_2, x_1)) = \frac{1}{1+e^{-(w_2 x_2 + w_1 x_1 + b)}} = \frac{1}{1+e^{-(w^T x + b)}}$$

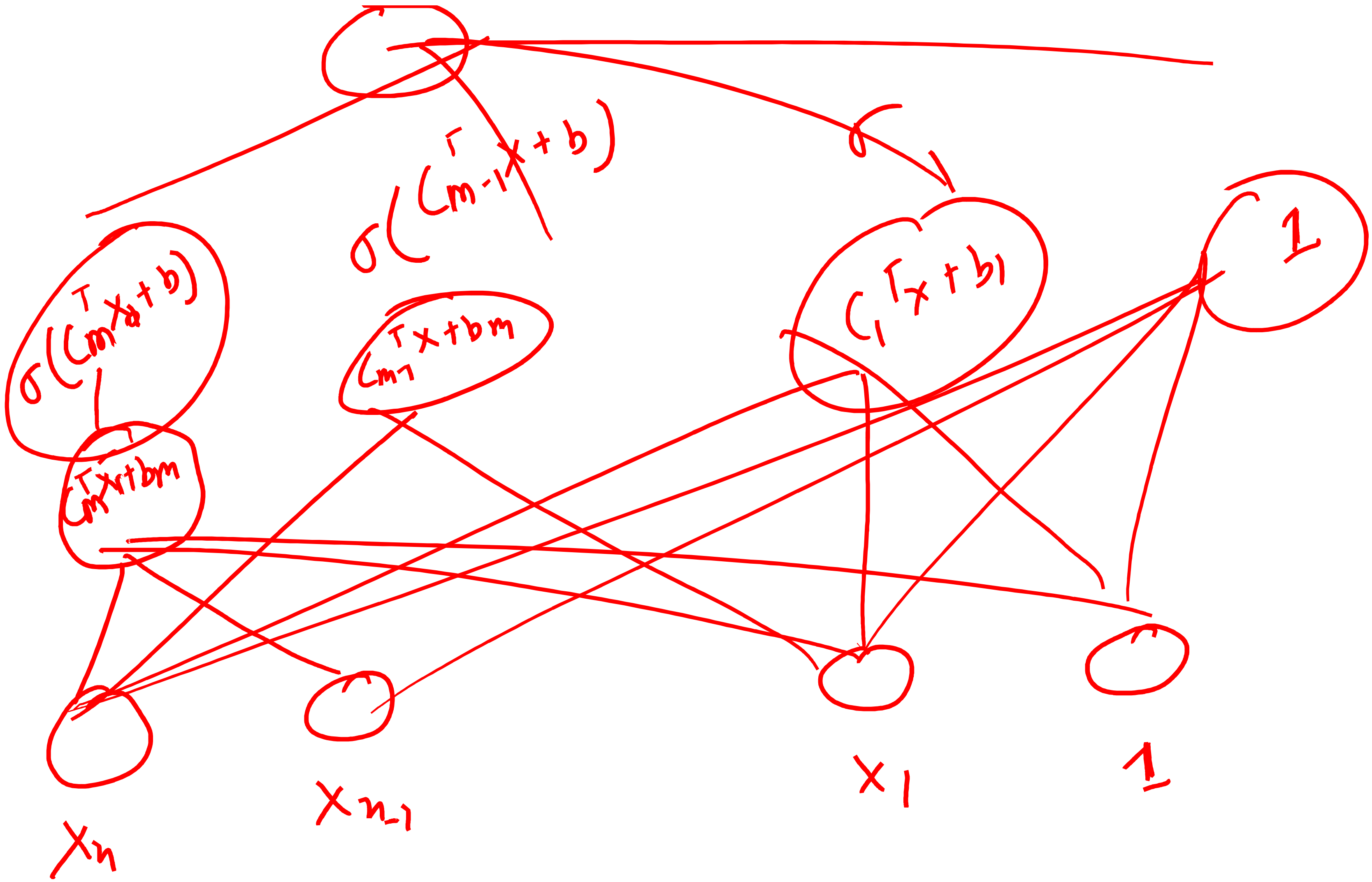
$$\sigma(x, w, b) = \frac{1}{1+e^{-(w_n x_n + w_{n-1} x_{n-1} + \dots + b)}} = \frac{1}{1+e^{-(w^T x + b)}}$$

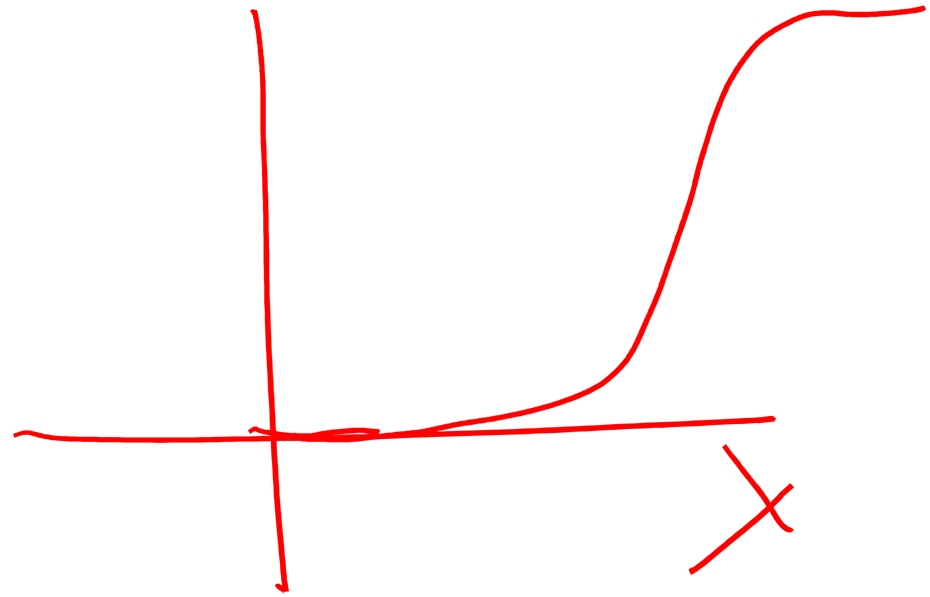
$$\sigma(x, c_1, b_1) = \frac{1}{1 + e^{-(c_1^T x + b_1)}} \quad \varphi_1(x)$$

$$\sigma(x, c_2, b_2) = \frac{1}{1 + e^{-(c_2^T x + b_2)}} \quad \varphi_2(x)$$

$$\sigma(x, c_m, b_m) = \frac{1}{1 + e^{-(c_m^T x + b)}} \quad \varphi_m(x)$$

$$= \underline{w^T \varphi(x) + b}$$


$\sigma(C_m^T x + b)$
$\sigma(C_{m-1}^T x + b)$
σ
$C_m^T x + b_m$
$C_{m-1}^T x + b_m$
$C_1^T x + b_1$
1
x_n
x_{n-1}
x_1
1


X