https://edintelligence.github.io/

ED INTELLIGENCE

We are looking for an enthusiastic new committee to run this year's EdIntelligence!

Facebook Event: https://bit.ly/2lWEuRe
Committee Interest Form: https://bit.ly/2mjwHNX

Are you interested in AI, Machine Learning and Deep Learning and want to make it more accessible to everyone?

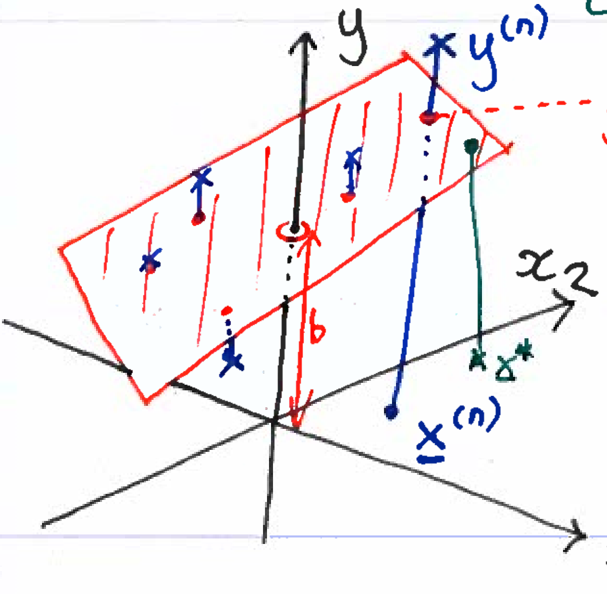
Our events are such as:

- Cutting Edge Research Talks by Researchers in Academia and the Industry
- Hackathons/Workshops in Machine Learning
- Recruiting events such as our annual AI Career Fair and recruitment talks with industry throughout the year
- Socials :)
- and much more

Sounds like something you want to be involved in? Then come to this meeting and figure out the specifics! Ask questions that you might have and nominate yourself for one of our committee roles from president to Career Fair team!

MLPR Lecture 3

{tinyurl.com/edmlpr}



$$f(\underline{x}^{(n)}; \underline{w}, b) = \underline{w}^\top \underline{x}^{(n)} + b = \underline{\tilde{w}}^\top \underline{\phi}(\underline{x}^{(n)})$$

$$\underline{\tilde{w}} = \begin{bmatrix} \underline{w} \\ b \end{bmatrix}, \quad \underline{\phi}(\underline{x}) = \begin{bmatrix} \underline{x} \\ 1 \end{bmatrix}$$

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

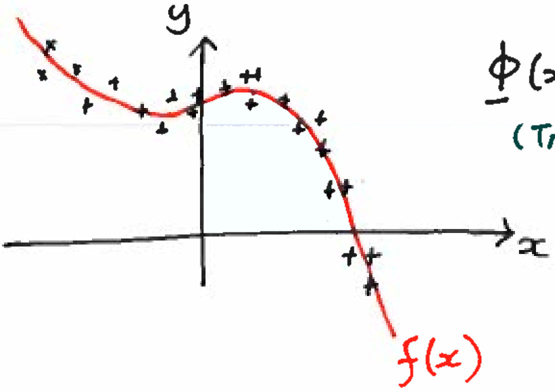
n^{th} row $\underline{\phi}(\underline{x}^{(n)})^\top$

$f_n = f(\underline{x}^{(n)}; \underline{v})$

choose $\underline{\tilde{w}}$ to minimize

$$(\underline{y} - \underline{f})^\top (\underline{y} - \underline{f})$$

" $\underline{\tilde{w}} = \Phi \backslash \underline{y}$ "



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

(Transforms 1D data $\rightarrow$ 4D data)

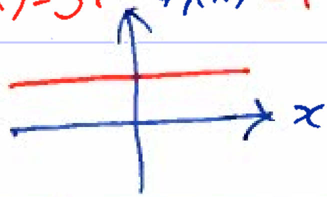
Fit

$y \approx \underline{f} = \overset{N \times 4}{\Phi}\,\overset{4 \times 1}{\underline{w}}$

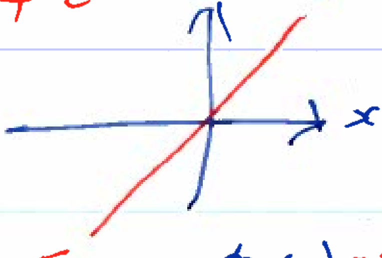
$f(x) = w_1 \cdot 1 + w_2 x + w_3 x^2 + w_4 x^3$

Basis functions

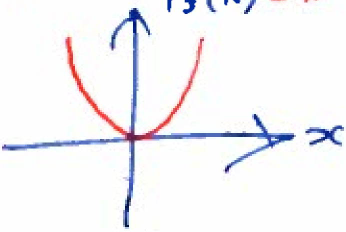
$f(x) = 5 \cdot$ $\phi_1(x) = 1$



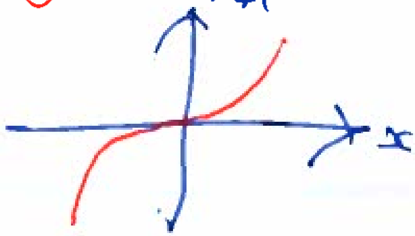
$+ 2 \cdot$ $\phi_2(x) = x$

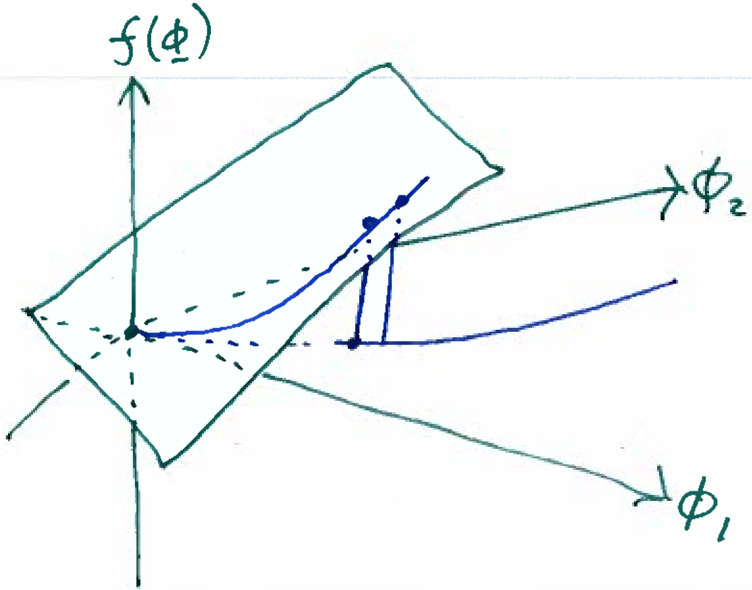


$-1 \cdot$ $\phi_3(x) = x^2$



$-5 \cdot$ $\phi_4(x) = x^3$



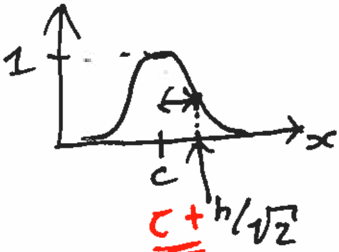


$$\underline{\Phi}(x) = \begin{bmatrix} x \\ x^2 \end{bmatrix}$$

$f(\underline{\Phi})$ is linear in $\underline{\Phi}$

$f(\underline{\Phi}(x))$ is non-linear in x

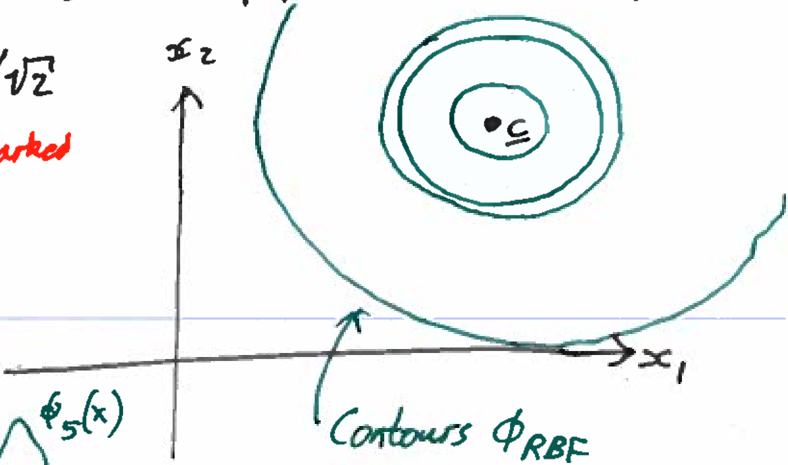
Radial Basis Function (RBF)

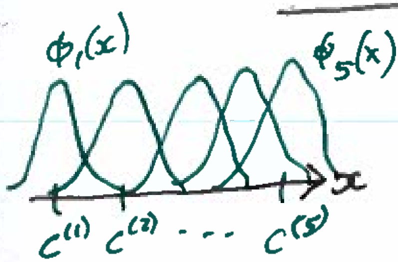


The width marked ↔ is $h/\sqrt{2}$

$$\phi_{RBF}(\underline{x};\, \underline{c}, h)$$

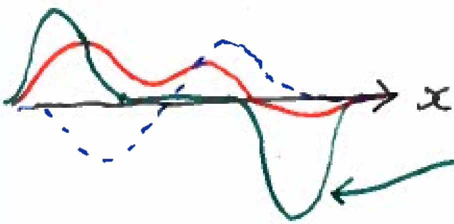
$$= \exp\left(-(\underline{x}-\underline{c})^T(\underline{x}-\underline{c})/h^2\right)$$





$$f(x) = \sum_{k=1}^{5} w_k \phi_k(x)$$

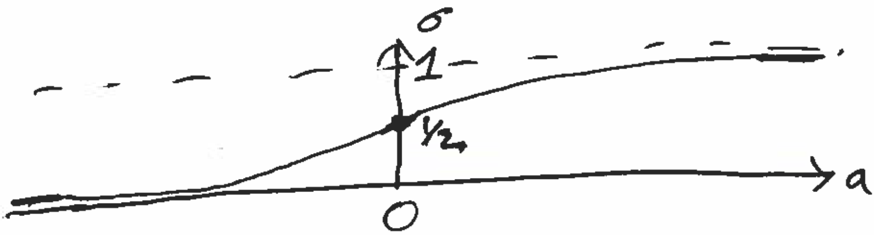
$$= \underline{w}^T \underline{\phi}(\underline{x})$$



$\underline{w} = [1\ 0\ 0\ 0\ -1]^T$

Logistic - Sigmoid function

$$\sigma(a) = \frac{1}{1+e^{-a}}$$



Basis f^n

$$\phi_\sigma(\underline{x}; \underline{v}, b) = \sigma(\underline{v}^T\underline{x} + b)$$

To do yourself: 2D contour plot.

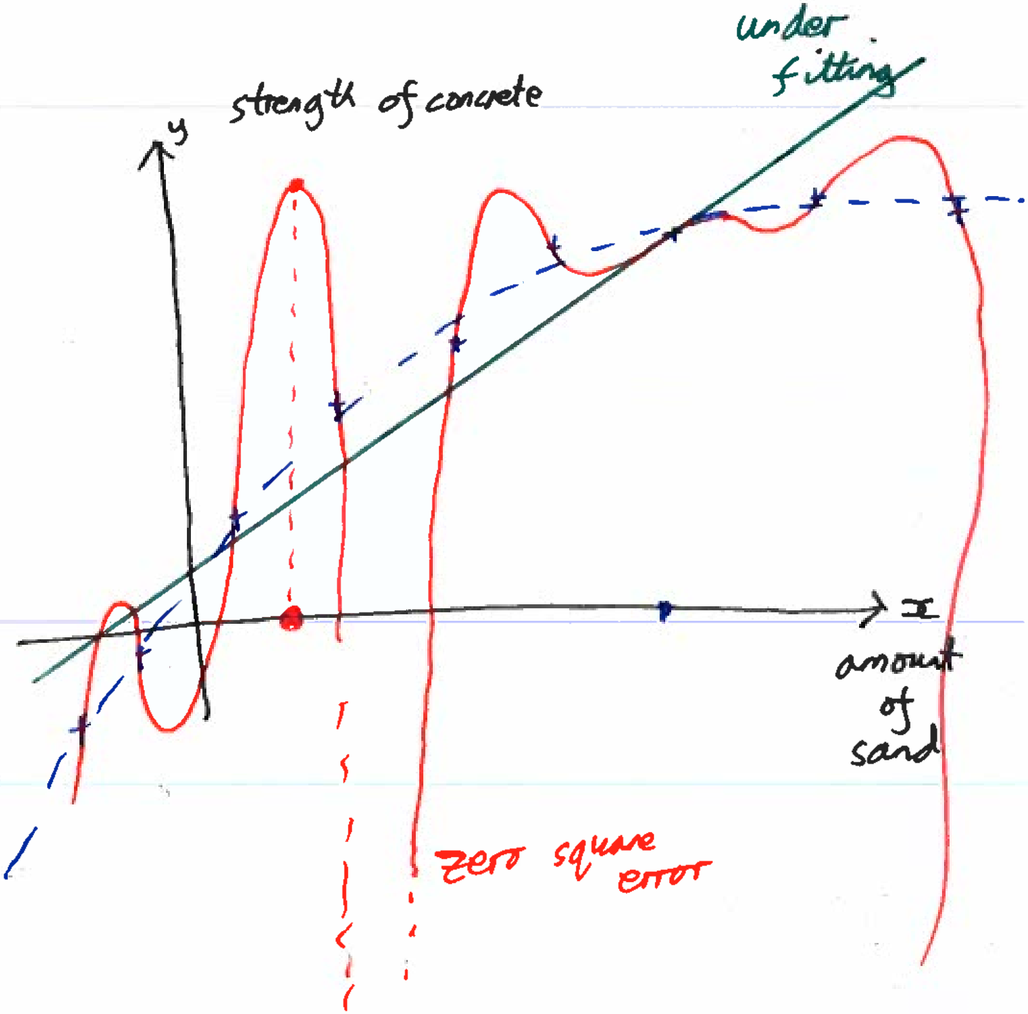
High-dimensional Polynomials

$$\underline{\phi}(\underline{x}) = [\ 1 \quad x_1 \quad x_2 \quad x_3 \ \ldots$$
$$x_1^2 \quad x_2^2 \quad x_3^2 \ \ldots$$
$$x_1 x_2 \quad x_1 x_3 \quad x_2 x_3 \ \ldots$$
$$x_1^3 \quad x_1 x_2 x_3 \quad x_1 x_2^2 \ \ldots$$
$$\ldots\ ]$$

Binary vectors can work well

XOR problem

x_1	x_2	$f(\underline{x})$
0	0	0
1	0	1
0	1	1
1	1	0



L2 Regularization

Discourage extreme fits

$\underline{w}^T\underline{w} = \|\underline{w}\|^2$ should be small

Cost function, which we minimize

$$C_\lambda(\underline{w}) = (\underline{y} - \underbrace{\Phi \underline{w}}_{\underline{f}})^T (\underline{y} - \Phi \underline{w}) + \lambda\, \underline{w}^T \underline{w}$$

$\lambda \in [0, \infty]$

$$\tilde{\underline{y}} = \left[\begin{array}{c} \underline{y} \\ \underline{0} \end{array}\right] \begin{array}{l} N \times 1 \\ K \times 1 \end{array} \qquad \tilde{\Phi} = \begin{array}{c} \\ N+K \end{array} \left\{ \underbrace{\left[\begin{array}{c} \Phi \\ \sqrt{\lambda}\, I_K \end{array}\right]}_{K} \right. \begin{array}{l} N \times K \\ \text{\# basis fn's} \end{array}$$

Fit $\quad \hat{\underline{w}} = \tilde{\Phi} \backslash \tilde{\underline{y}}$

Minimizes

$$\hat{\underline{w}} = \underset{\underline{w}}{\operatorname{argmin}} (\tilde{\underline{y}} - \tilde{\Phi}\underline{w})^T (\tilde{\underline{y}} - \tilde{\Phi}\underline{w})$$