2018 L4 ①

tinyurl.com/edmlpr

(Try refreshing the page)

Tutorials:
- 1st sheet up
- Meetings next week (TBA soon!)
- Answers released end next week

Assignment pairs:

See website

Hypothesis Forum

- Share links, code snippets
- Get code review
- Ask Q's
- Post answers < help others / get feedback

Linear Regression Reminders

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

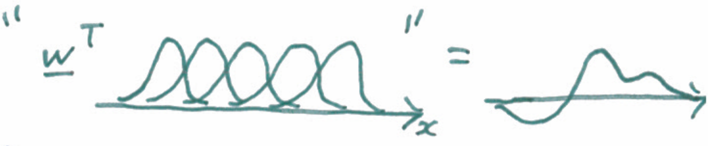
Can minimize

$$\underbrace{\textstyle\sum_n \left(y^{(n)} - \underline{w}^T \underline{\phi}(\underline{x}^{(n)})\right)^2}_{= (y - \Phi \underline{w})^T (y - \Phi \underline{w})} \quad \text{wrt } \underline{w}$$

$$\underline{\phi}(\underline{x}) = \left[\phi_1(\underline{x}) \;\; \phi_2(\underline{x}) \;\; \ldots \;\; \phi_K(\underline{x})\right]^T$$

$\phi_k(\underline{x})$ any scalar function

- Monomial, eg x_2, $x_3 x_4^3$, ...
- RBF



- Sigmoid

Why are large weights bad?

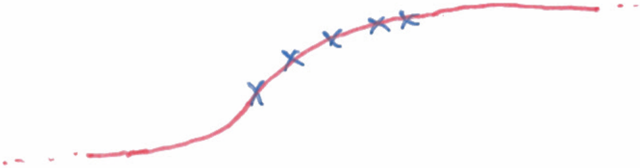
If $\underline{w}$ are bounded then function bounded

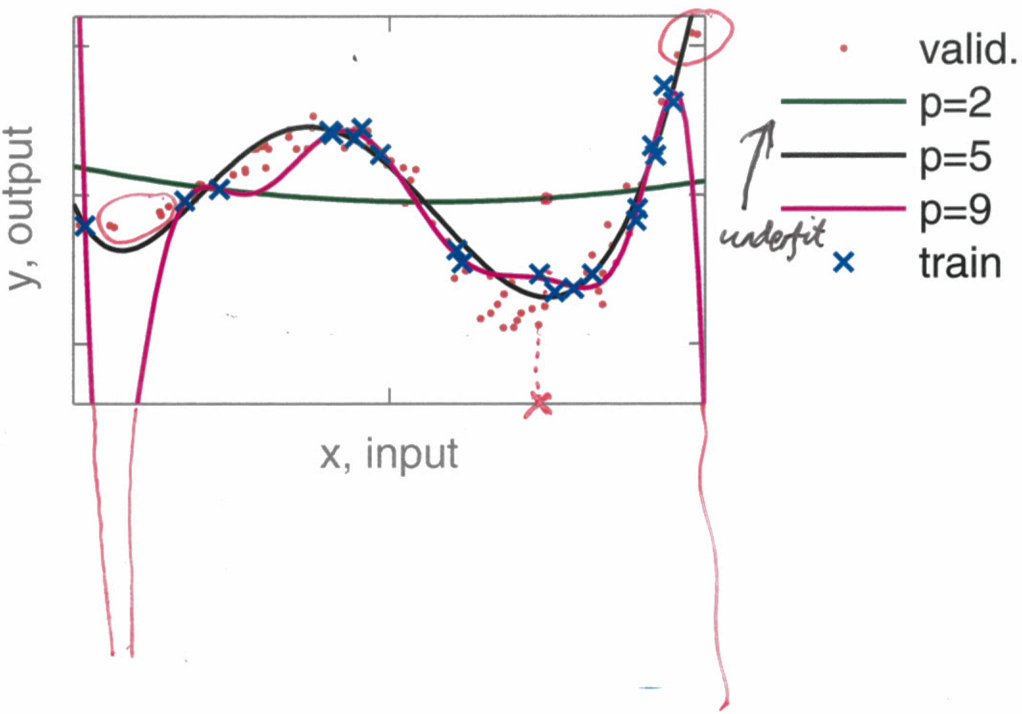
Large derivatives are bad.

If $\underline{w}$ are bounded $\rightarrow$ derivatives are bounded.

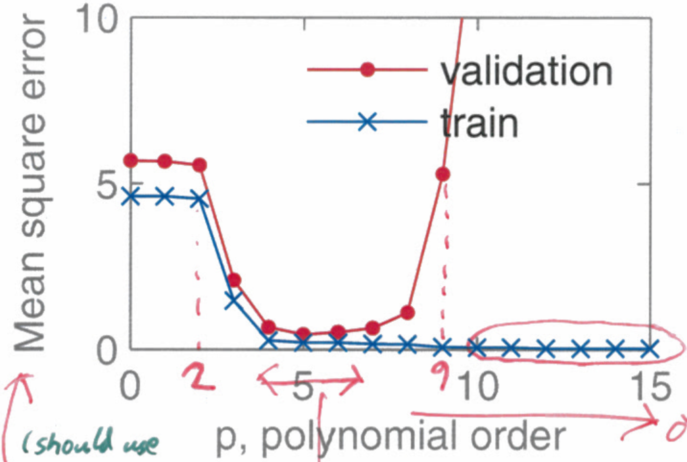
RBF: Always extrapolates to 0

$\sigma(x)$





2018 L4 4



overfitting

(should use log scale)

reasonable

underfitting

regularization λ

Generalization

$$E_{gen} = \mathbb{E}_{p(\underline{x},y)}\left[L(y,\ f(\underline{x})) \right]$$

Loss function

Assuming there is some fixed distribution on future inputs & outputs

$$= \iint L(y, f(\underline{x}))\, p(\underline{x}, y)\, d\underline{x}\, dy$$

Monte Carlo unbiased approximation

$$\approx \frac{1}{M} \sum_{m=1}^{M} L(y^{(m)}, f(\underline{x}^{(m)})) = E_{test}$$

$$y^{(m)}, \underline{x}^{(m)} \sim p(\underline{x}, y)$$

Draw samples from a held out test set.

Data Splits

Training set: fit $\underline{w}$

Don't fit:

Order of a polynomial

\# of RBFs

Regularization constant, λ

Validation Set

To fit λ, model choices

Test set

To report estimate of generalization error.

Reading: kaggle blog.