tinyurl.com/edmlpr

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

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

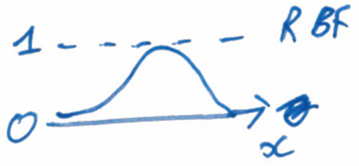
$$= (y - \Phi \underline{w})^T (y - \Phi \underline{w})$$

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

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

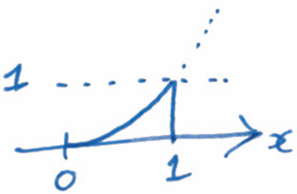
- Monomial, eg x_2, $x_3 x_4^3$, ...
- RBF

- Sigmoid

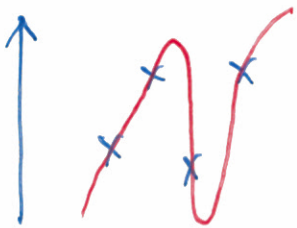


If $\underline{w}$ are bounded

then $f^\wedge$ bounded



(Chebfun)



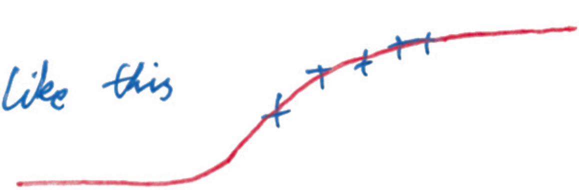
Large derivatives are bad

If $\underline{w}$ are bounded

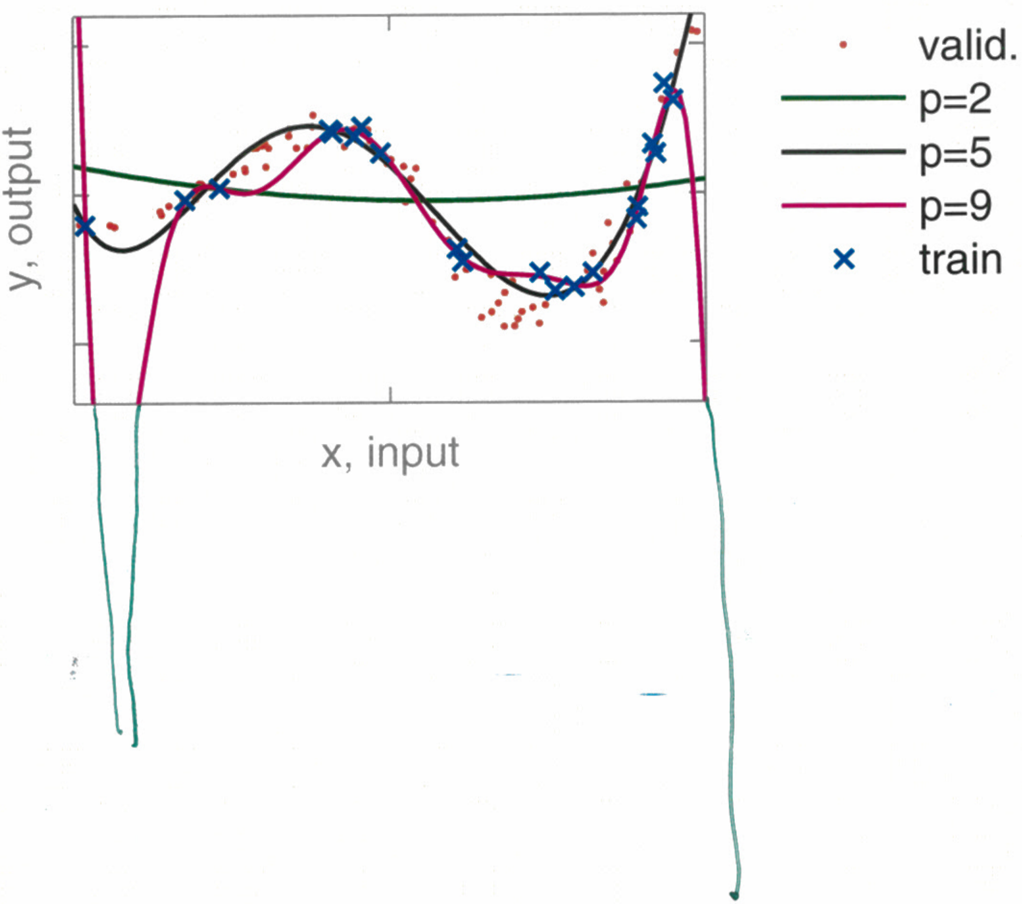
→ derivatives also bounded.

RBF always extrapolates to 0.

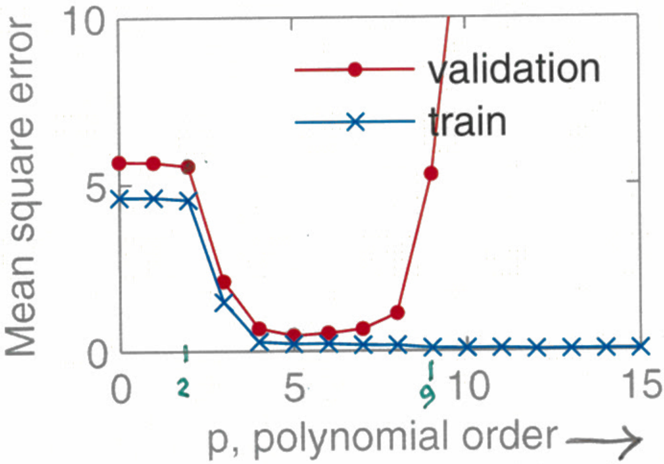
Sigmoids extrapolate like this



2017
L4 4



2014
L4 5



← log λ regularization const

Generalization

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

Loss function

We assume there is some fixed distribution $p(\underline{x}, y)$ on future inputs & outputs

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

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 examples from held out test set.

But not if model was selected so E_{test} is small.

Data Splits

Training set: fit $\underline{w}$

(Don't fit:
Order of a polynomial
# of RBFs
Regularization constants.)

Validation set:
(Development set)
To fit λ, model choices

Test set: To report estimate
of generalization error.

Reading: kaggle blog.