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

Can minimize \[ \sum_n (y^{(n)} - w^T \phi(x^{(n)}))^2 \]

wrt \( w \) \[
= (y - \Phi w)^T (y - \Phi w)
\]

\( \phi(x) = [\phi_1(x), \phi_2(x), \ldots, \phi_k(x)]^T \)

\( \phi_k(x) \) any scalar function
- Monomial, e.g. \( x_2, x_3 x_4^3, \ldots \)
- RBF
- Sigmoid

\[ w^T \rightarrow \sigma(x) \]
If \( w \) are bounded
then \( f \) is bounded

(Chebfun)

Large derivatives are bad
If \( w \) are bounded
\( \rightarrow \) derivatives also bounded.

RBF always extrapolates to 0.

Sigmoid extrapolates like this.
Mean square error

Validation
Train

$p$, polynomial order

$log$ regularization constant
Generalization

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

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

\[ E_{\text{gen}} = \iint L(y, f(x)) p(x, y) \, dx \, dy \]

unbiased

Monte Carlo approximation

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

\[ y^{(m)}, x^{(m)} \sim p(x, y) \]

Draw examples from held out test set.

But not if model was selected so \( E_{\text{test}} \) is small.
Data Splits

Training set: fit w

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

Validation set: (Development set)

To fit \( \lambda \), model choices

Test set: To report estimate of generalization error.

Reading: Kaggle blog.