

# Least Angle Regression

Tim Hesterberg, Insightful Corp.

16 June 2006

This is joint work with Chris Fraley, with support from NIH SBIR  
Phase I 1 R43 GM074313-01

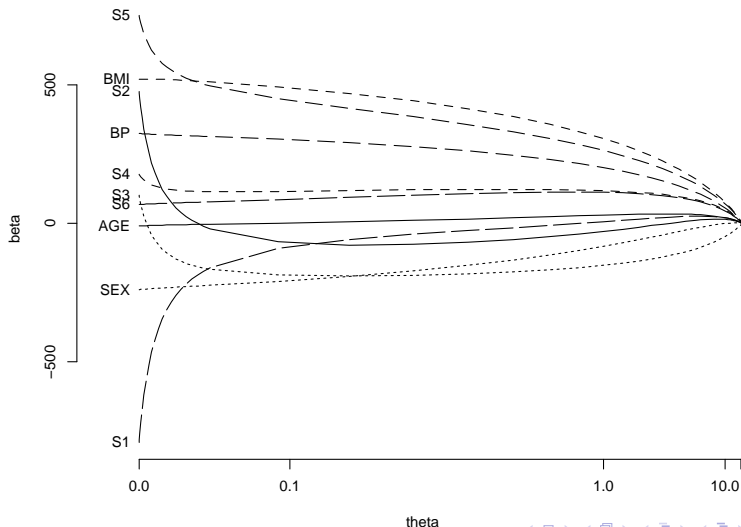
- ▶ Why is LARS important?
- ▶ Other packages
- ▶ GLARS package
- ▶ Issues
- ▶ Insightful Research

# Why is LARS important?

- ▶ Variable Selection in Regression
  - ▶ Important
  - ▶ Many approaches: stagewise, boosting, LASSO, regularization, ...
- ▶ Least Angle Regression — Efron, Hastie, Johnstone, Tibshirani (2004) *Annals* (with discussion)
  1. Lasso
  2. Forward stagewise
  3. Least Angle Regression (LAR)
    - ▶ Unifying explanation
    - ▶ Fast implementation
    - ▶ Fast way to choose tuning parameter

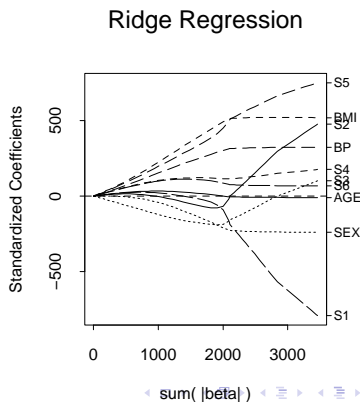
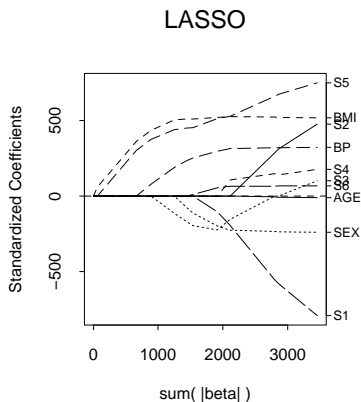
# Ridge Regression

- ▶ Minimize  $\sum(Y_i - \hat{Y}_i) + \lambda \sum \hat{\beta}_j^2$



# LASSO

- ▶ Minimize  $\sum(Y_i - \hat{Y}_i) + \lambda \sum |\hat{\beta}_j|$
- ▶ Forces small coefficients  $\rightarrow 0$ ; gives simpler models.
- ▶ Smaller penalty on large coefficients: less effect on important terms
- ▶ Implementation is more complicated and slower



# Forward Stagewise Regression

(Forward Stagewise = Least Squares Boosting)

1. Initialize: standardize predictors, center  $y$ ,  
 $r = y, \beta_1 = \dots = \beta_p = 0$
2. Repeat many times
  - ▶ Find the predictor  $x_j$  most correlated with  $r$
  - ▶  $\delta = \epsilon \text{sign}(r \cdot x_j)$
  - ▶  $\hat{\beta}_j \leftarrow \hat{\beta}_j + \delta$
  - ▶  $r \leftarrow r - \delta x_j$

# Forward Stagewise and LASSO

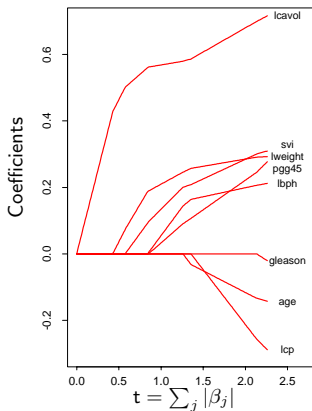
March 2003

Trevor Hastie, Stanford Statistics

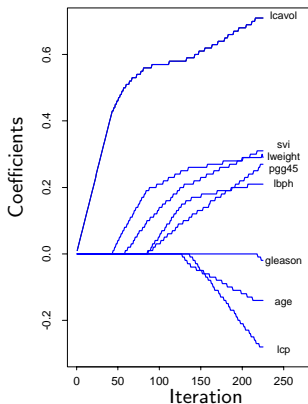
6

## Prostate Cancer Data

Lasso



Forward Stagewise



# Similarity:

Are LASSO and infinitesimal forward stagewise identical?

- ▶ With orthogonal predictors, yes.
- ▶ Otherwise similar.

Least Angle Regression provides explanation, and fast implementation.

# Stepwise, Forward Stagewise, Least Angle

## Stepwise regression:

- ▶ Pick predictor most correlated with  $y$
- ▶ Bring predictor completely into model (full LS fit)

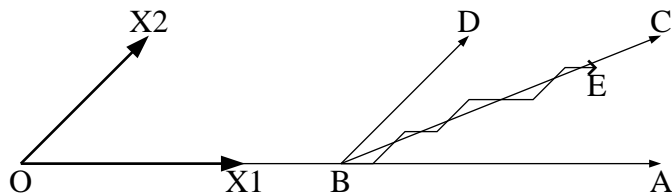
## Forward stagewise:

- ▶ Pick predictor most correlated with  $y$
- ▶ Increment coefficient for predictor

## Least Angle Regression:

- ▶ Pick predictor most correlated with  $y$
- ▶ Bring predictor into model only to extent it is better than others
- ▶ Move in least-squares direction until another variable is as correlated

# Least Angle Regression



$C$  = projection of  $y$  onto space spanned by  $X_1$  and  $X_2$ .

$B$  = first step for least-angle regression

$E$  = point on stagewise path

# LARS - other packages

`lars` : Efron and Hastie (S-PLUS and R)

- ▶ Linear regression

`glmpath` : Park and Hastie (R)

- ▶ GLM and Cox Proportional Hazards

Methods: `plot`, `print`, `predict`, `cv`, `coef`

- ▶ S-PLUS and R, open source
  - ▶ Incorporate `lars`, `glm`path
  - ▶ Cleanup, consistent interface
  - ▶ Incorporate future work by others; provide framework
- ▶ Extensions
  - ▶ Numerically-accurate calculations
  - ▶ Factors, splines, polynomials, interactions, ...
  - ▶ Other models (robust regression, ...), other penalties
  - ▶ Missing data
  - ▶ Massive data sets
  - ▶ Diagnostics, tools for selecting tuning parameter
- ▶ User-friendly
  - ▶ Consistent interface
  - ▶ GUI
  - ▶ Documentation

- ▶ Money
  - ▶ NIH funding: require commercial potential
  - ▶ Insightful: indirect benefit
- ▶ Outside contributors
- ▶ Licensing; ability to ship with S-PLUS, I-Miner.

# Insightful Research Department

- ▶ Turn research into software for wide use
  - ▶ Higher standards than academic software (ease of use, robustness, testing)
- ▶ Collaboration
- ▶ Variety: resampling, missing data, group sequential designs, simulation-based econometric software, functional data, stable distributions, proteomics, microarrays, frailty models, causal modeling
- ▶ External funding — SBIR grants (NIH, NSF, ...)
  - ▶ Somewhat easier funding
  - ▶ Commercial potential
  - ▶ Risk, research element
- ▶ We're hiring
- ▶ We're looking for good projects and collaborators