

STANFORD UNIVERSITY  
DEPARTMENT OF STATISTICS  
DEPARTMENTAL SEMINAR

4:15 p.m., Tuesday, January 8, 2002  
Sequoia Hall Room 200  
(Cookies at 3:45 in 1st Floor Lounge)

*Bradley Efron*  
*Department of Statistics*  
*Stanford University*

**Least Angle Regression**

Automatic model-building algorithms are familiar, or notorious, in the linear model literature: forward stepwise, backward stepwise, all subsets, and various combinations, are used to produce "good" linear models for predicting a response  $y$  on the basis of some measured quantities  $x[1], x[2], \dots, x[p]$ . Goodness is defined in terms of prediction accuracy, but also in terms of parsimony: simpler models are preferred for the sake of scientific insight into the  $x$ - $y$  relationship.

The "Lasso" and "Forward Stagewise Linear Regression" are promising new model-building algorithms that work differently than the traditional methods. They give results that are strikingly similar but not identical to each other, as shown in the examples in the recent book by Hastie, Tibshirani, and Friedman. It turns out that both are variants of a simpler technique, "Least Angle Regression". Besides helping understand the Lasso/Stagewise methodology, the Least Angle Regression algorithm is computationally much more efficient. It is also easier to analyze in terms of its statistical properties.

Joint work with Trevor Hastie and Rob Tibshirani.