# Regression

Greg Grudic



# Admin

- If you haven't, please send in your survey.

## Your Concerns

1. Math Skills
2. Matlab
3. Workload



## Matlab

- The link to the matlab page is: http://www.mathworks.com/academia/student_version/
- Almost all of the ITS lab machines have matlab

## Questions?



## To Do…

- Homework 1 is on the class web page
- Please read next weeks lecture notes before next class…

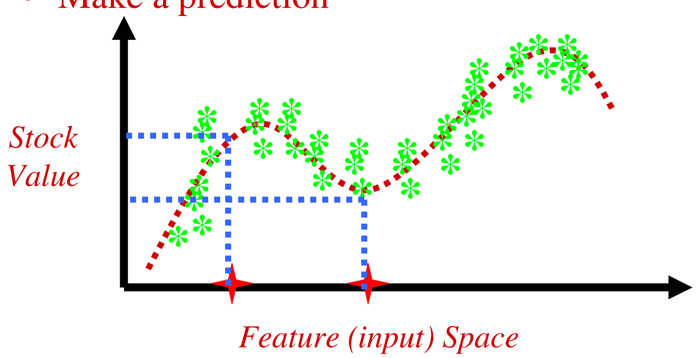
## Lecture Goal

- Basic definitions
- Introduction to Regression
  - Least squares regression
  - Ridge Regression
  - Lasso Regression
  - Kernels
  - Kernel Ridge Regression

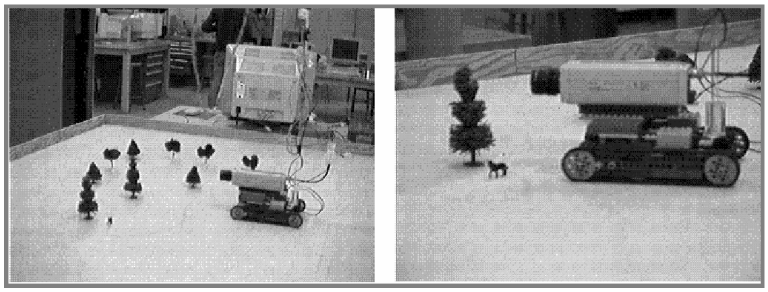


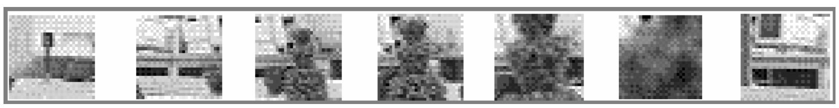
## Learning Regression Models

- Collect Training data
- Build Model: stock value = F(feature space)
- Make a prediction



## My Favorite Regression Example:







# Human-to-Robot Skill Transfer (ICRA96)

- **Problem:** Human demonstrates a task via teleoperation.
  - Object locate and approach task.
  - 1024 raw pixel inputs and 2 actuator outputs.
- **Learning Data:** 4 demonstrations of task sequence.
  - 2000 to 5000 learning examples (~2 to 5 min).
- **Learning Time:** ~5 min. on SPARC 20.
- **Model Size / Evaluation Speed:** < 500 Kb, ~5 Hz.
- **Autonomous control of robot using model:**
  - No failures in 30 random trials.

## Ridge Regression

Given learning data:

$$\{(\mathbf{x}_1, y_1),...,(\mathbf{x}_N, y_N)\}, \text{ where } \mathbf{x}_i \in \Re^d, y_i \in \Re$$

Define: $$\mathbf{G} = \begin{pmatrix} g_{11} & \cdots & g_{1K} \\ \vdots & \ddots & \vdots \\ g_{N1} & \cdots & g_{NK} \end{pmatrix} \text{ and } \overline{g}_i = \frac{1}{N}\sum_{a=1}^{N} g_{ai}$$

Where:

1. Linear Ridge: $g_{ij} = x_{ij}$ and $K = d$
2. Kernel Ridge: $g_{ij} = K(\mathbf{x}_i, \mathbf{x}_j)$ and $K = N$

Kernel Matrix



## Ridge Coefficients

$$\mathbf{X} = \begin{pmatrix} g_{11} - \overline{g}_1 & \cdots & g_{1K} - \overline{g}_K \\ \vdots & \ddots & \vdots \\ g_{N1} - \overline{g}_1 & \cdots & g_{NK} - \overline{g}_K \end{pmatrix} = \mathbf{G} - \begin{pmatrix} \overline{g}_1 & \cdots & \overline{g}_K \\ \vdots & \ddots & \vdots \\ \overline{g}_1 & \cdots & \overline{g}_K \end{pmatrix}$$

$$\left(\hat{\beta}_1^{ridge},...,\hat{\beta}_K^{ridge}\right)^T = \left(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I}\right)^{-1}\mathbf{X}^T\mathbf{y}$$

$$\hat{\beta}_0^{ridge} = \left[\frac{1}{N}\sum_{i=1}^{N} y_i\right] - \left[\sum_{j=1}^{K} \overline{g}_j \hat{\beta}_j^{ridge}\right]$$

# Learned Models

- Linear

$$\hat{y} = \hat{\beta}_0 + \sum_{j=1}^{d} \hat{\beta}_j x_j$$

- Kernel

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^{N} \hat{\beta}_i K\left(\mathbf{x}_i, \mathbf{x}\right)$$