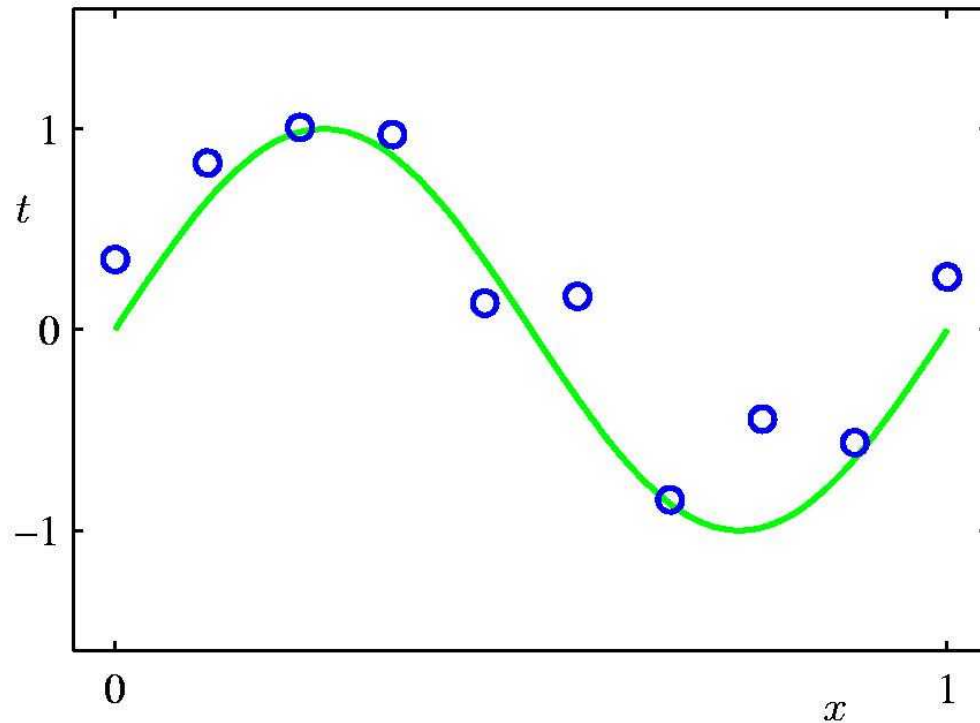

Linear Algebra

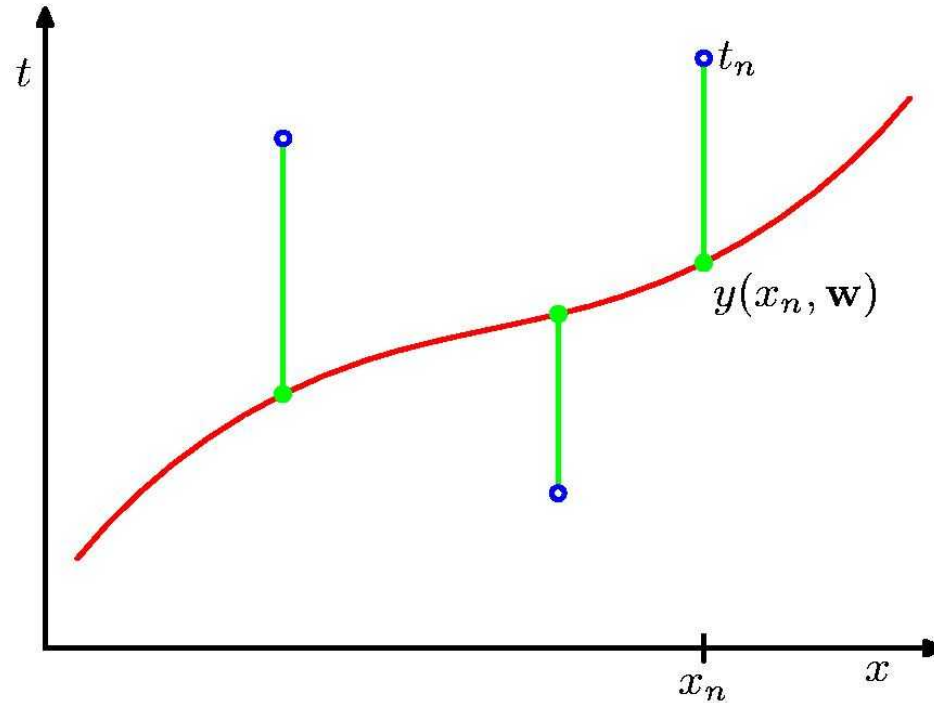
Curve Fitting

Last Class: Curve Fitting



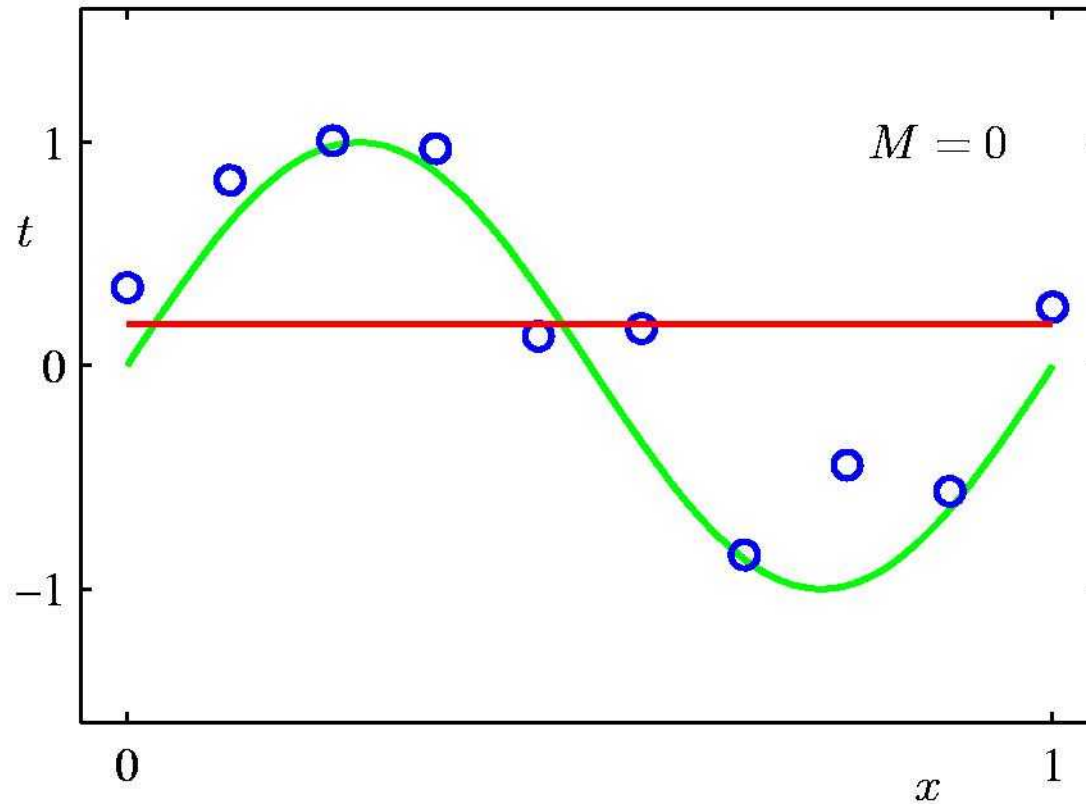
$$y(x, \mathbf{w}) = w_0 + w_1x + w_2x^2 + \dots + w_Mx^M = \sum_{j=0}^M w_jx^j$$

Sum-of-Squares Error Function

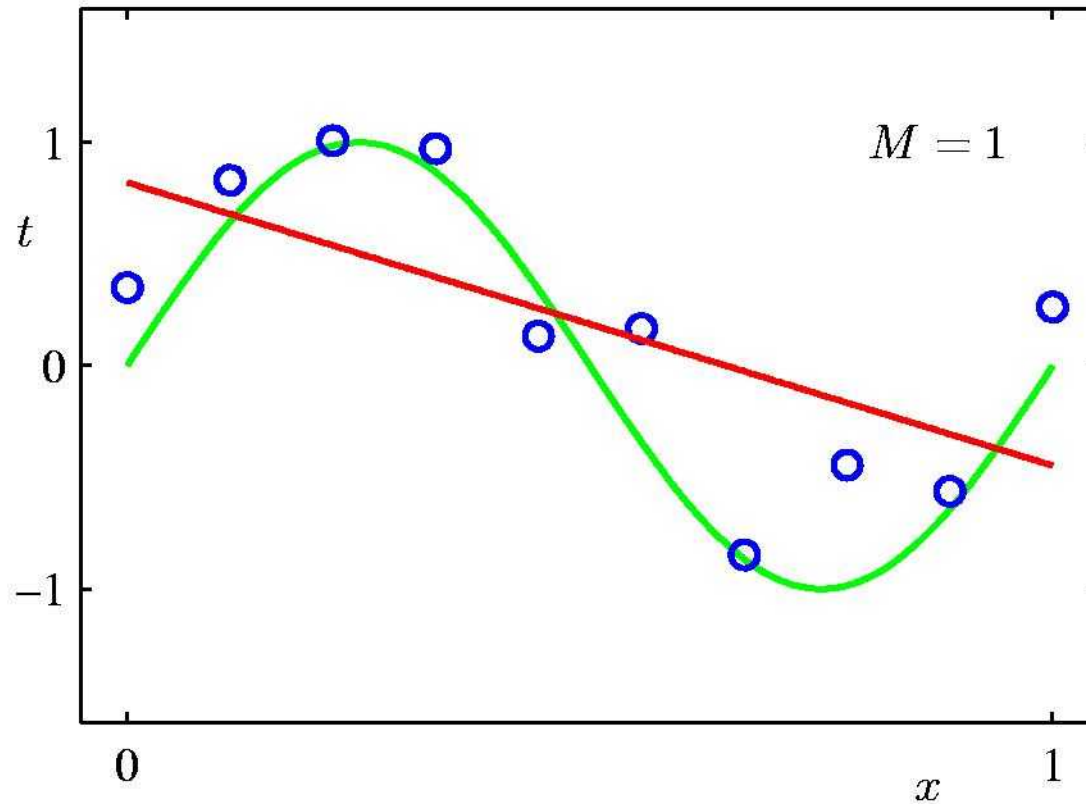


$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2$$

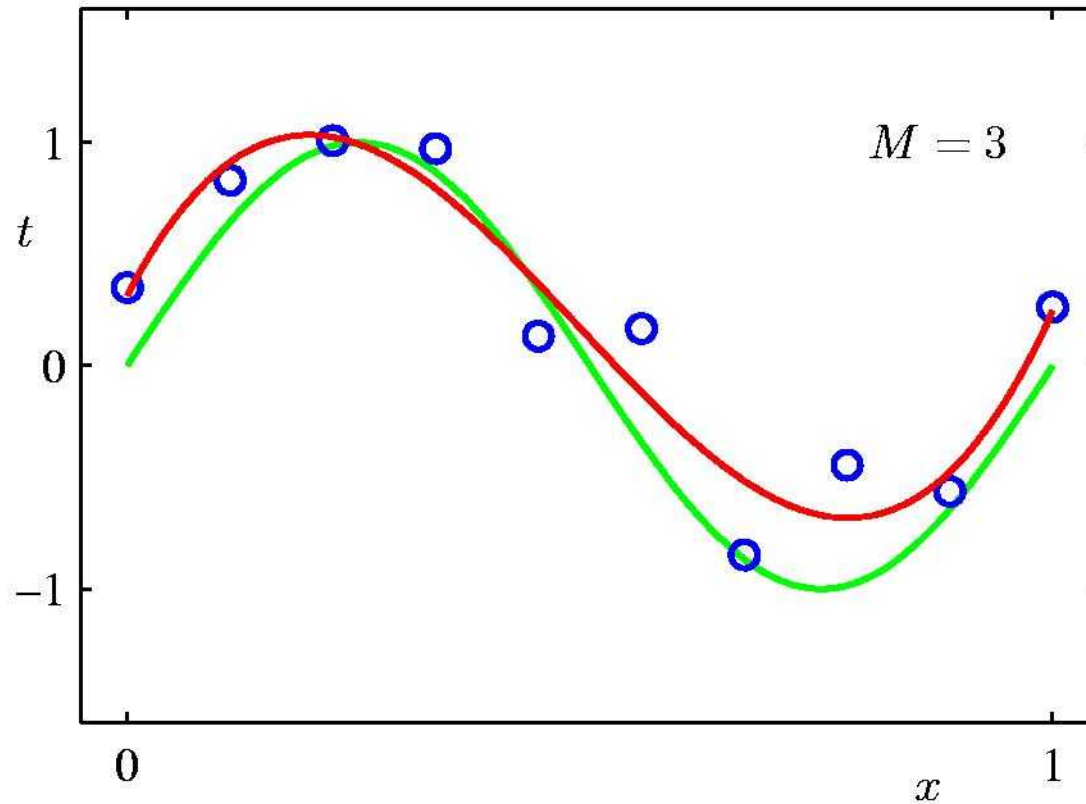
0th Order Polynomial



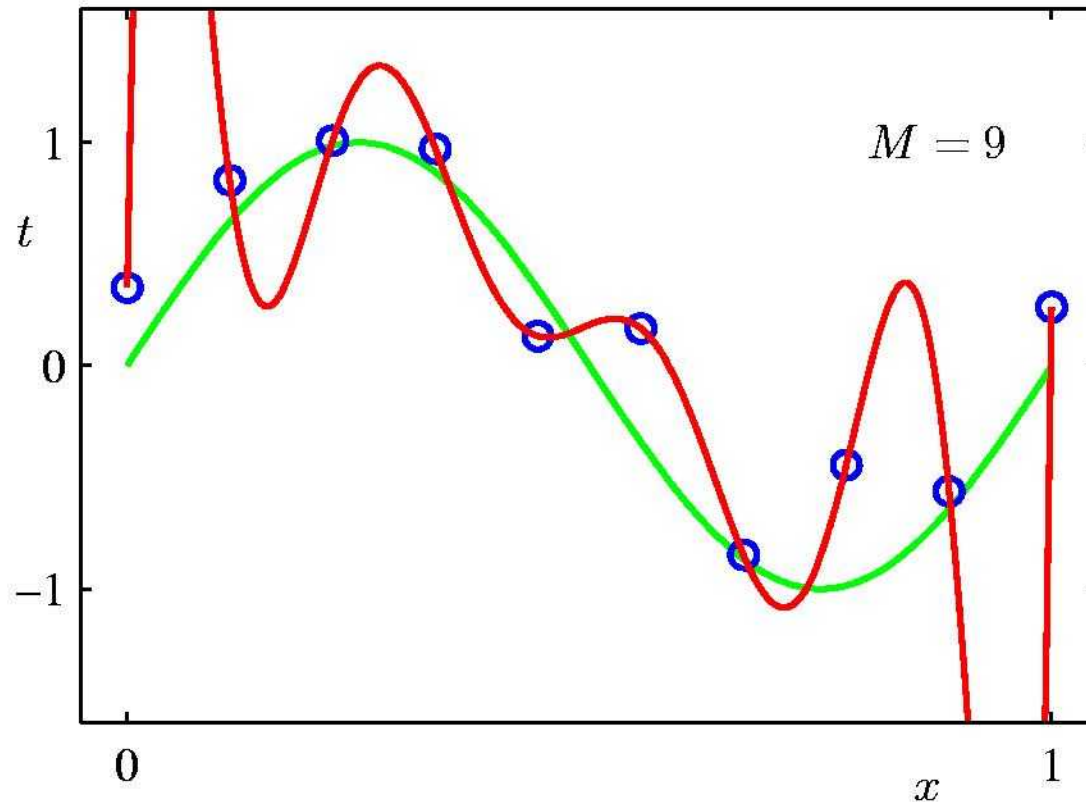
1st Order Polynomial



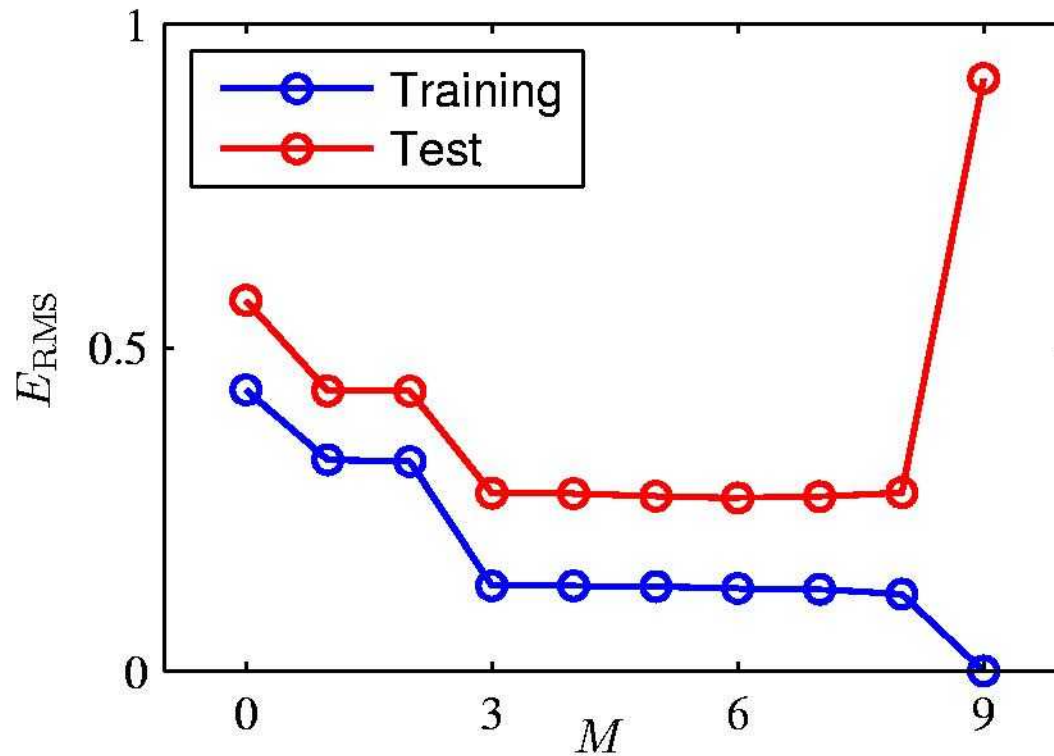
3rd Order Polynomial



9th Order Polynomial



Over-fitting



Root-Mean-Square (RMS)

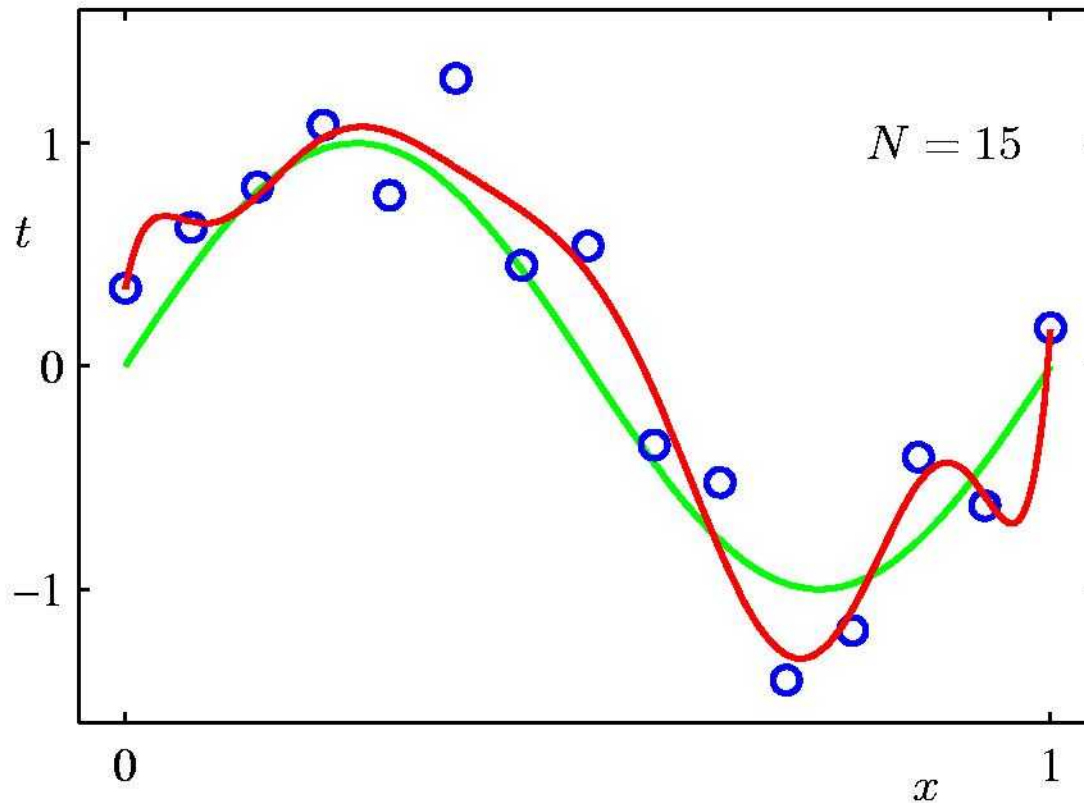
Error: $E_{\text{RMS}} = \sqrt{2E(\mathbf{w}^*)/N}$

Polynomial Coefficients

| | $M = 0$ | $M = 1$ | $M = 3$ | $M = 9$ |
|---------|---------|---------|---------|-------------|
| w_0^* | 0.19 | 0.82 | 0.31 | 0.35 |
| w_1^* | | -1.27 | 7.99 | 232.37 |
| w_2^* | | | -25.43 | -5321.83 |
| w_3^* | | | 17.37 | 48568.31 |
| w_4^* | | | | -231639.30 |
| w_5^* | | | | 640042.26 |
| w_6^* | | | | -1061800.52 |
| w_7^* | | | | 1042400.18 |
| w_8^* | | | | -557682.99 |
| w_9^* | | | | 125201.43 |

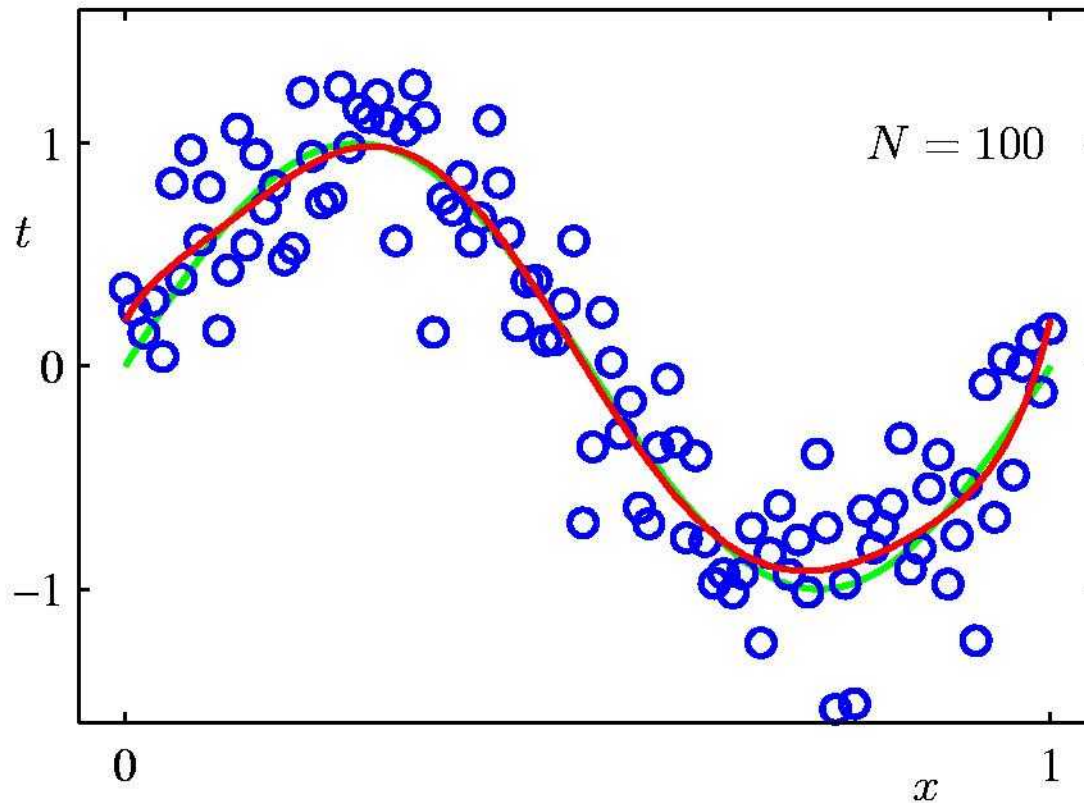
Data Set Size: $N = 15$

9th Order Polynomial



Data Set Size: $N = 100$

9th Order Polynomial

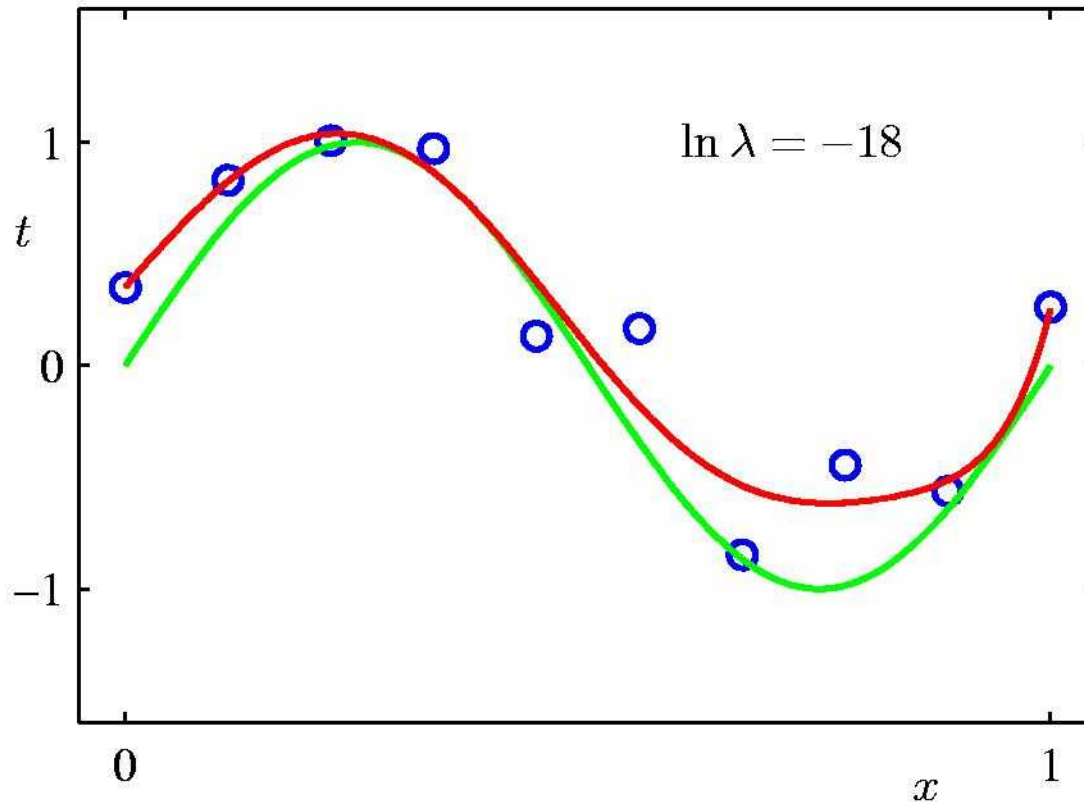


Regularization

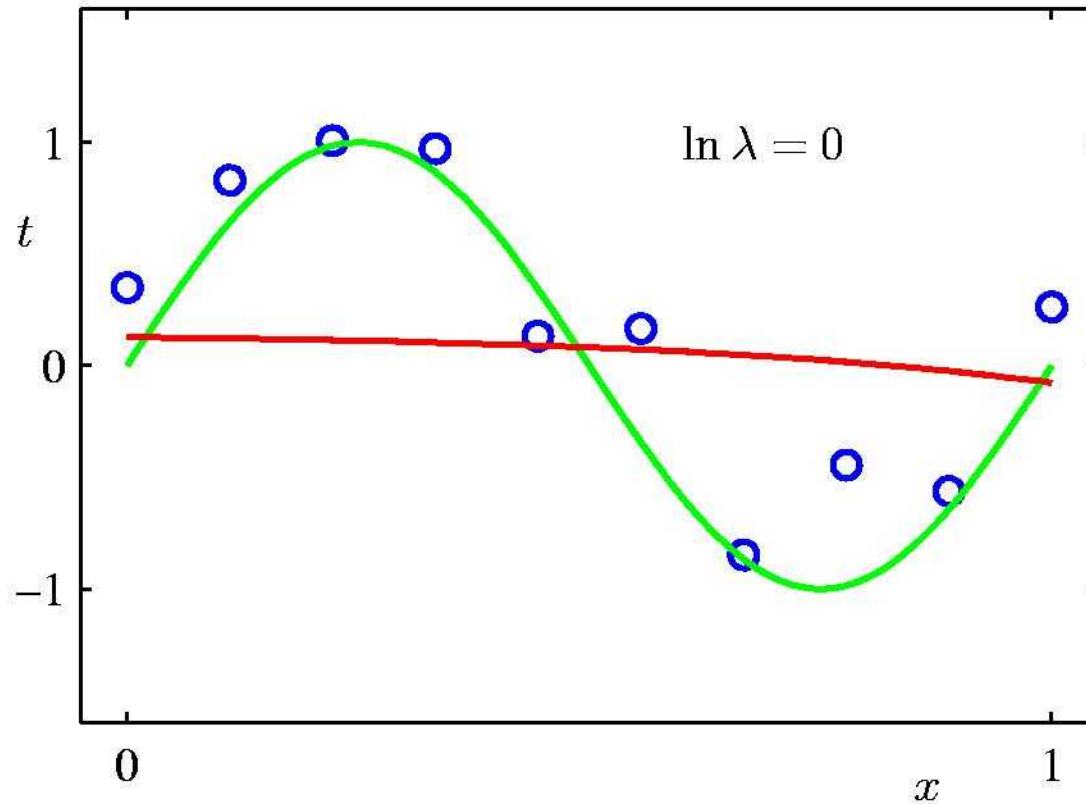
Penalize large coefficient values

$$\tilde{E}(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2 + \frac{\lambda}{2} \|\mathbf{w}\|^2$$

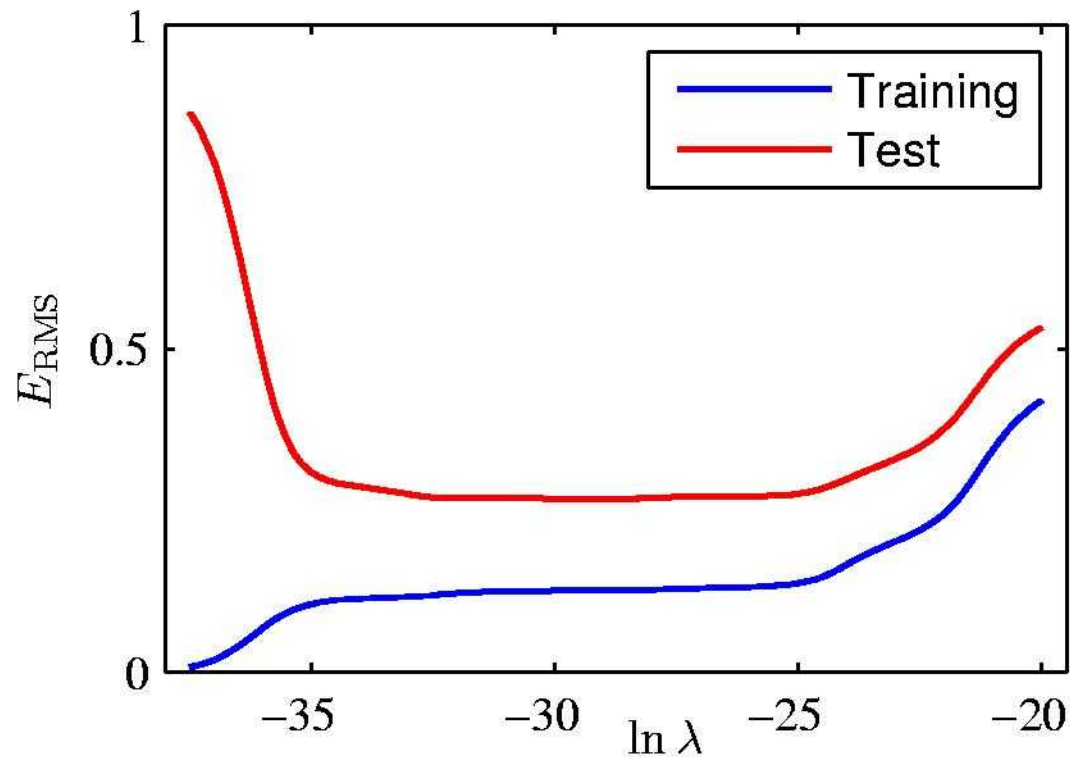
Regularization: $\ln \lambda = -18$



Regularization: $\ln \lambda = 0$



Regularization: E_{RMS} vs. $\ln \lambda$



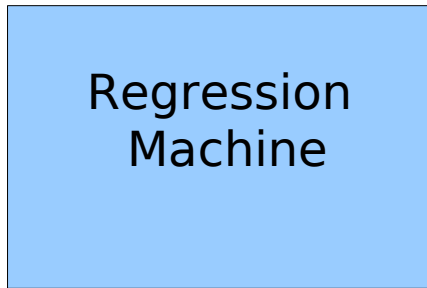
Polynomial Coefficients

| | $\ln \lambda = -\infty$ | $\ln \lambda = -18$ | $\ln \lambda = 0$ |
|---------|-------------------------|---------------------|-------------------|
| w_0^* | 0.35 | 0.35 | 0.13 |
| w_1^* | 232.37 | 4.74 | -0.05 |
| w_2^* | -5321.83 | -0.77 | -0.06 |
| w_3^* | 48568.31 | -31.97 | -0.05 |
| w_4^* | -231639.30 | -3.89 | -0.03 |
| w_5^* | 640042.26 | 55.28 | -0.02 |
| w_6^* | -1061800.52 | 41.32 | -0.01 |
| w_7^* | 1042400.18 | -45.95 | -0.00 |
| w_8^* | -557682.99 | -91.53 | 0.00 |
| w_9^* | 125201.43 | 72.68 | 0.01 |

Methodology

| x | y |
|-----|----|
| 0.5 | 10 |
| 0.6 | 1 |
| 0.8 | 12 |

↓ Training

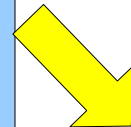
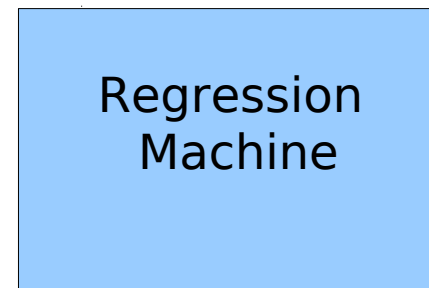
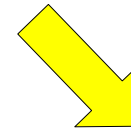


Update weights w and degree M s.t. min error

$$\tilde{E}(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2 + \frac{\lambda}{2} \|\mathbf{w}\|^2$$

Regression

1



0.5

Questions

- **How to** find the curve that best fits the points?

i.e., how to solve the minimization problem?

$$\tilde{E}(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2 + \frac{\lambda}{2} \|\mathbf{w}\|^2$$

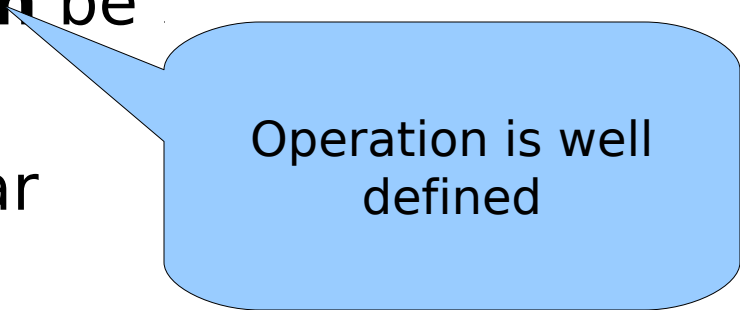
- **Answer 1:** use gnuplot or any other pre-built software (homework)
 - **Answer 2:** recall linear algebra (will see with more details in upcoming class)
-
- **Why** solve this particular minimization problem?

e.g., why solve this minimization problem rather than doing linear interpolation?

Principled answer follows...

Linear Algebra

- **To find curve that best fits points:** apply basic concepts from linear algebra
- **What is linear algebra?**
 - The study of vector spaces
- **What are vector spaces?**
 - Set of elements that **can** be
 - Summed
 - Multiplied by scalar



Operation is well defined

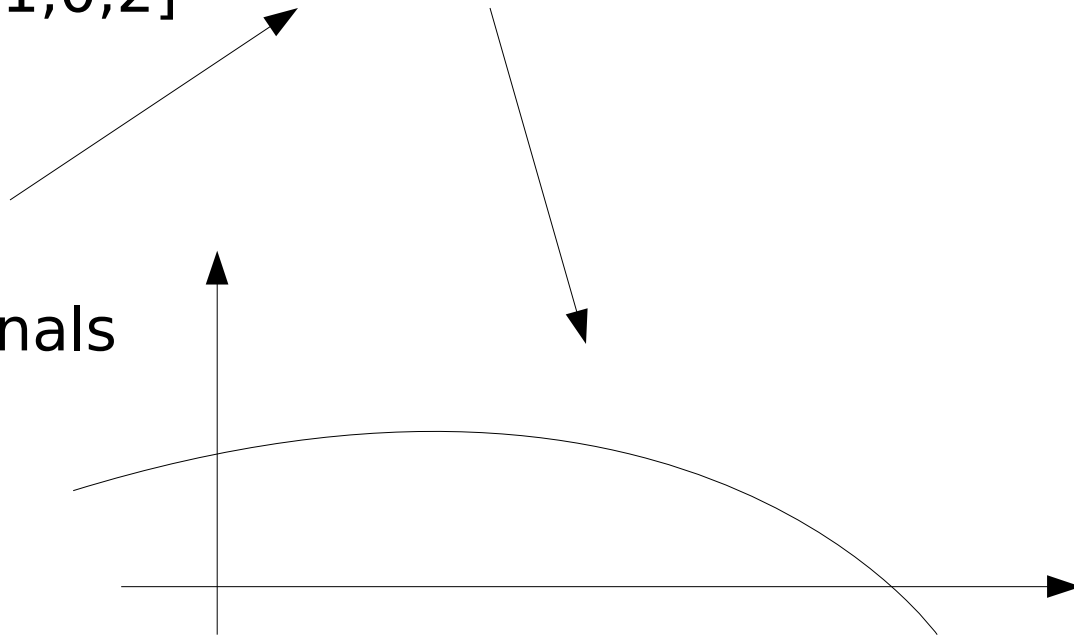
Linear Algebra=Study of Vectors

- Examples of vectors

- $[0,1,0,2]$

-

- Signals



- **Signals are vectors!**

Dimension of Vector Space

- **Basis = set of element that span space**
 - Dimension of line = 1
 - basis = $(1,0)$
 - Dimension of plane = 2
 - basis = $\{(1,0), (0,1)\}$
 - Dimension of space = 3
 - basis = $\{(1,0,0), (0,1,0), (0,0,1)\}$
-
- **Dimension of Space = # Elements Basis**
-

Functions are Vectors I

• **Vectors**

- **Size**

- $x=[x_1,x_2]$

- $|x|^2=x_1^2+x_2^2$

- **Inner product**

- $y=[y_1,y_2]$

- $x \cdot y = x_1 y_1 + x_2 y_2$

Functions

Size

$$y=f(t)$$

$$|y|^2 = \int f^2(t) dt$$

Inner Product

$$z=g(t)$$

$$yz = \int f(t) g(t) dt$$

Signals are Vectors II

• **Vectors**

- $x=[x_1,x_2]$
- Basis
 - $\{[1,0],[1,0]\}$
- Components
 - x_1
 - x_2

Functions

$$y=f(t)$$

Basis

$\{1, \sin \omega t, \cos \omega t, \sin 2\omega t, \cos 2\omega t, \dots\}$

Components

coefficients of basis functions

<http://www.falstad.com/fourier/>

A Very Special Basis

• Vectors

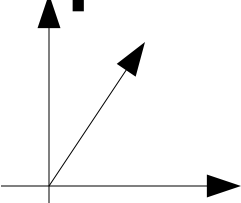
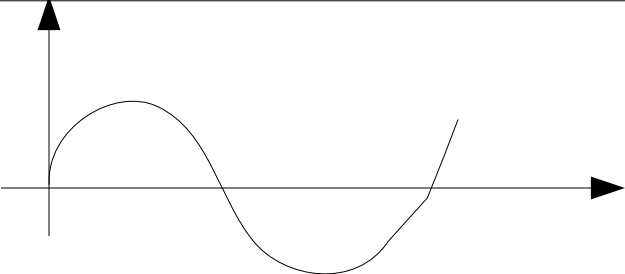
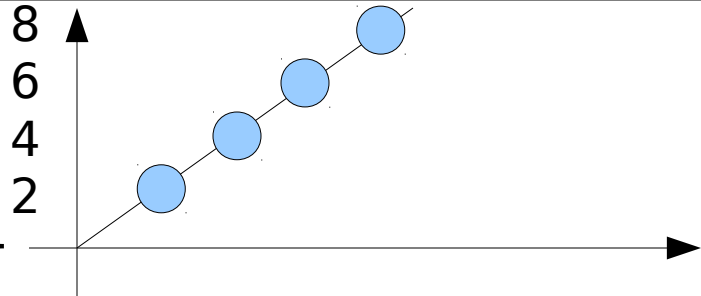
- $x = [x_1, x_2]$
- Basis
 - $\{[1, 0]$
- Components
 - x_1
 - x_2

Basis in which
each element
is a vector,
formed by
set of M samples
of a function
(Subspace of \mathbb{R}^M)

Functions

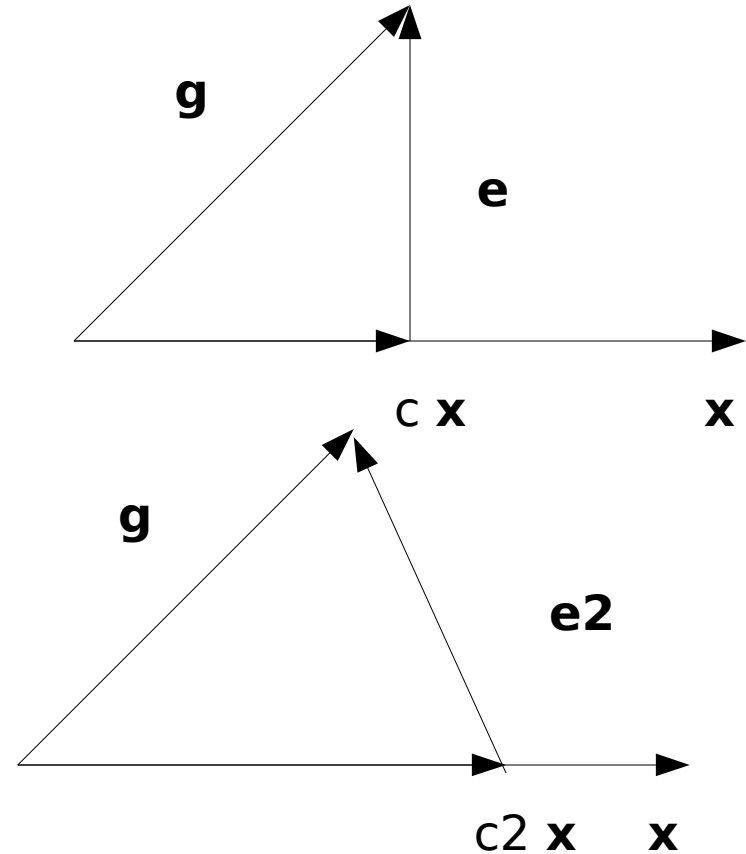
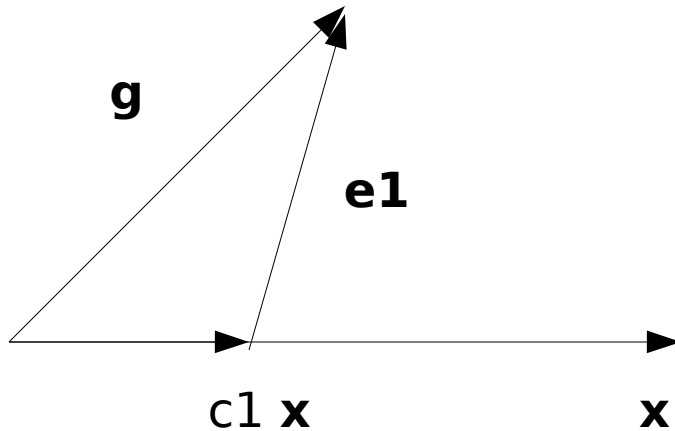
- $y = f(t)$
- Basis
- $\{1, \sin \omega t, \cos \omega t, \sin 2\omega t, \cos 2\omega t, \dots\}$
- Components
- coefficients of basis
functions

Representations

| | Coeff. | Basis | Representation |
|---------------------|---------------|--|--|
| traditional | [1,2] | [0,1] [1,0] |  |
| function | [0,1] | {1, sin(wt), ... } |  |
| sampled function | [0,2,0] | {[1,1,1,1], [1,2,3,4], [1,4,9,16]} |  |

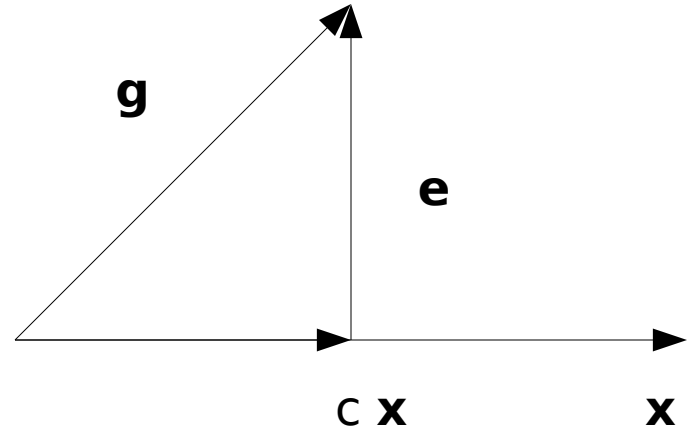
Approximating a Vector I

- $\mathbf{e} = \mathbf{g} - c \mathbf{x}$
- $\mathbf{g}' \mathbf{x} = |\mathbf{g}| |\mathbf{x}| \cos \varphi$
- $|\mathbf{x}|^2 = \mathbf{x}' \mathbf{x}$
- $|\mathbf{g}| \cos \varphi = c |\mathbf{x}|$



Approximating a Vector II

- $\mathbf{e} = \mathbf{g} - c \mathbf{x}$
- $\mathbf{g}' \mathbf{x} = |\mathbf{g}| |\mathbf{x}| \cos \varphi$
- $|\mathbf{x}|^2 = \mathbf{x}' \mathbf{x}$
- $|\mathbf{g}| \cos \varphi = c |\mathbf{x}|$

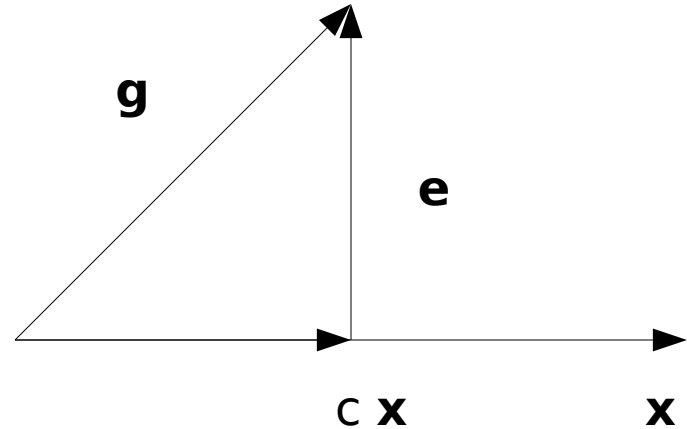


- **Major result: $c = \mathbf{g}' \mathbf{x} (\mathbf{x}' \mathbf{x})^{-1}$**

$$|e|^2 = (g_1 - cx_1)^2 + (g_2 - cx_2)^2 = \sum (g_i - cx_i)^2$$

Approximating a Vector III

- $\mathbf{e} = \mathbf{g} - c \mathbf{x}$
- $\mathbf{g}' \mathbf{x} = |\mathbf{g}| |\mathbf{x}| \cos \varphi$
- $|\mathbf{x}|^2 = \mathbf{x}' \mathbf{x}$
- $|\mathbf{g}| \cos \varphi = c |\mathbf{x}|$
- $c = \mathbf{g}' \mathbf{x} / |\mathbf{x}|^2$



- **Major result:** $c = \mathbf{g}' \mathbf{x} ((\mathbf{x}' \mathbf{x})^{-1})'$

Coefficient to compute

Given target point
(basis matrix coords)

Given basis matrix

Approximating a Vector IV

- $\mathbf{e} = \mathbf{g} - \mathbf{x} \mathbf{c}$
- $\mathbf{g}' \mathbf{x} = |\mathbf{g}| |\mathbf{x}| \cos \varphi$
- $|\mathbf{x}|^2 = \mathbf{x}' \mathbf{x}$
- $\mathbf{g}' \mathbf{x} = \mathbf{x}' \mathbf{c}$

\mathbf{c} =vector of computed coefficients
 \mathbf{g} =vector of target points
 \mathbf{x} =basis matrix
- each column is a basis elem.
- each column is a polynomial evaluated at desired points

- **Major result: $\mathbf{c} = \mathbf{g}' \mathbf{x} ((\mathbf{x}' \mathbf{x})^{-1})'$**

$$|e|^2 = \sum (g_i - c x_i)^2$$

i-th line of matrix \mathbf{x}

Approximating a Vector IV

- $\mathbf{e} = \mathbf{g} - \mathbf{x} \mathbf{c}$
- $\mathbf{g}' \mathbf{x} = |\mathbf{g}| |\mathbf{x}| \cos \varphi$
- $|\mathbf{x}|^2 = \mathbf{x}' \mathbf{x}$
- $\mathbf{g}' \mathbf{x} = \mathbf{x} \mathbf{c}$

\mathbf{c} =vector of computed coefficients
 \mathbf{g} =vector of target points
 \mathbf{x} =basis matrix
- each column is a basis elem.
- each column is a polynomial evaluated at desired points

- **Major result: $\mathbf{c} = \mathbf{g}' \mathbf{x} ((\mathbf{x}' \mathbf{x})^{-1})'$**

Coefficient to compute

Given target point
(vector space V)

Given basis matrix
(elements in
basis matrix are
in vector space U
contained in V)

Least Square Error Solution!

- **As most machine learning problems, least square is equivalent to basis conversion!**

c=vector of computed coefficients
g=vector of target points
x=basis matrix

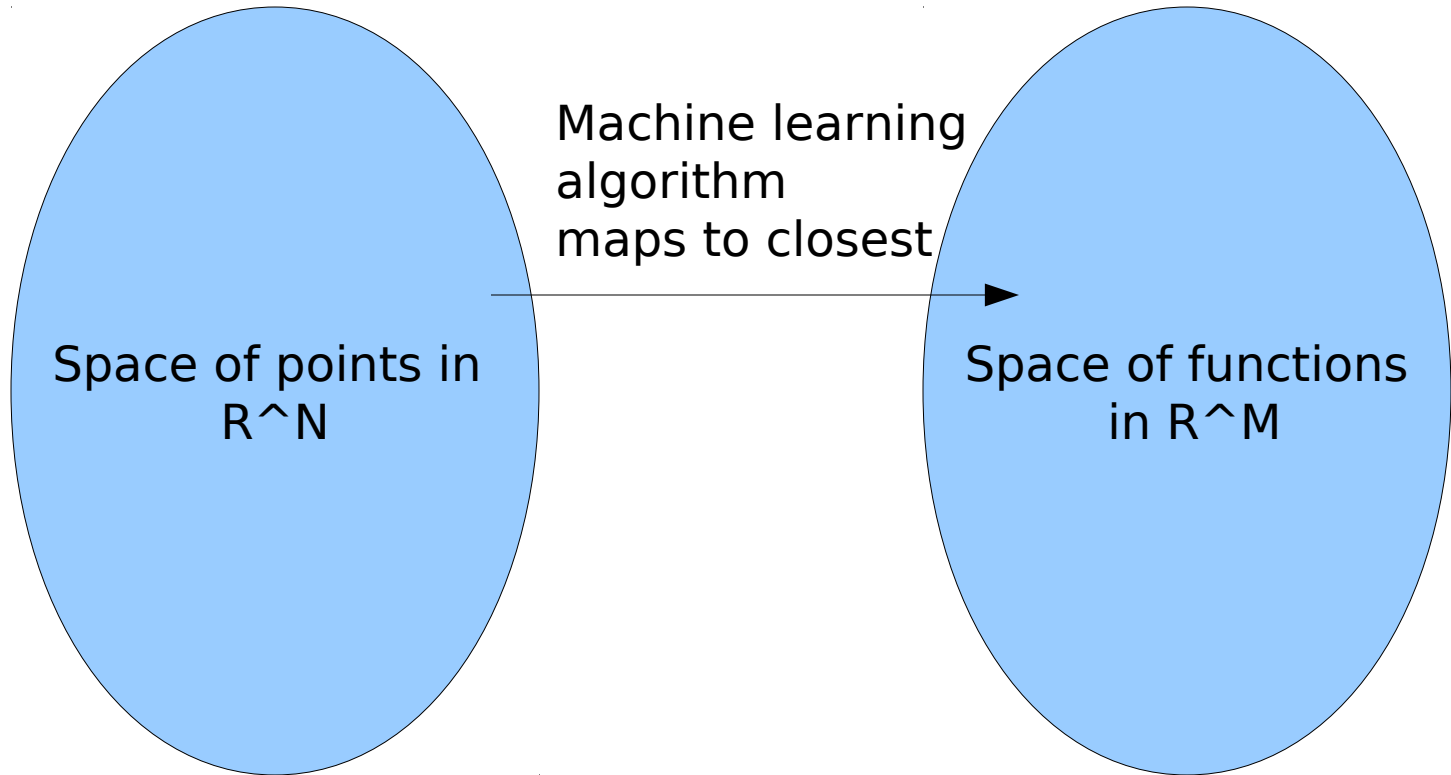
- each column is a basis elem.
- each column is a polynomial evaluated at desired points

- **Major result: $c = g' x ((x' x)^{-1})'$**

Coefficients to compute
(point in R^M)

Given points
(element in R^N)

Least Square Error Solution!



Regularized Least Square

c=vector of computed coefficients
g=vector of target points
x=basis matrix
- each column is a basis elem.
- each column is a polynomial
evaluated at desired points

- **Major result: $c = g' x ((\lambda I + x' x)^{-1})'$**
-

Least Squares and Gauss



- **How** to solve the least square problem?
 - Carl Friedrich **Gauss** solution in 1794 (age 18)
 - **Why** solving the least square problem?
 - Carl Friedrich **Gauss** solution in 1822 (age 46)
 - Least square solution is optimal in the sense that it is the best linear unbiased estimator of the coefficients of the polynomials
 - **Assumptions:** errors have zero mean and equal variances
- http://en.wikipedia.org/wiki/Least_squares
-