2.830J / 6.780J / ESD.63J Control of Manufacturing Processes (SMA 6303)
Spring 2008

Control of Manufacturing Processes

Subject 2.830/6.780/ESD.63

Spring 2008

Lecture #15

Response Surface Modeling and Process Optimization

April 8, 2008

Outline

- Last Time
  - – Fractional Factorial Designs
  - – Aliasing Patterns
  - – Implications for Model Construction

- Today
  - – Response Surface Modeling (RSM)
    - Regression analysis, confidence intervals
  - – Process Optimization using DOE and RSM

Reading: May & Spanos, Ch. 8.1 – 8.3

Regression Fundamentals

- Use least square error as measure of goodness to estimate coefficients in a model
- One parameter model:
  - Model form
  - Squared error
  - Estimation using normal equations
  - Estimate of experimental error
  - Precision of estimate: variance in *b*
  - Confidence interval for β
  - Analysis of variance: significance of *b*
  - Lack of fit vs. pure error
- Polynomial regression

Measures of Model Goodness – R^2

- Goodness of fit – R^2
  - Question considered: how much better does the model do than just using the grand average?

$$R^2 = \frac{SS_T}{SS_D}$$

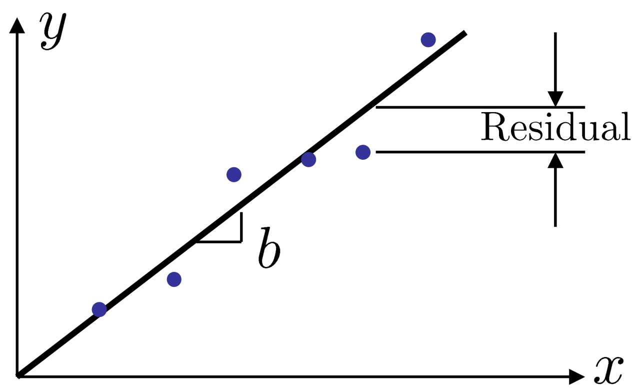
  - Think of this as the fraction of squared deviations (from the grand average) in the data which is captured by the model
- Adjusted R^2
  - For "fair" comparison between models with different numbers of coefficients, an alternative is often used

$$R^2_{\text{adj}} = 1 - \frac{SS_R/\nu_R}{SS_D/\nu_D} = 1 - \frac{s_R^2}{s_D^2}$$

  - Think of this as (1 – variance remaining in the residual). Recall $\nu_R = \nu_D - \nu_T$

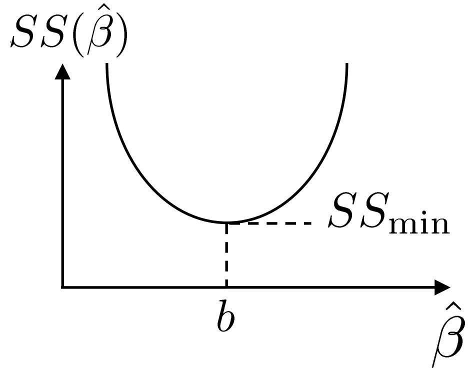
Least Squares Regression

- We use ***least-squares*** to estimate coefficients in typical regression models
- $$\begin{aligned} y_i &= \beta x_i + \epsilon_i, \quad i = 1, 2, ..., n; \;\; \epsilon_i \sim N(0, \sigma^2) \\ \hat{y}_i &= b x_i \end{aligned}$$



- Goal is to estimate β with "best" *b*
- How define "best"?
  - That b which minimizes sum of squared error between prediction and data

$$SS(\hat{\beta}) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - \hat{\beta} x_i)^2$$



  - The residual sum of squares (for the best estimate) is

$$SS_{\min} = \sum_{i=1}^{n} (y_i - b x_i)^2 = SS_R$$

Least Squares Regression, cont.

- Least squares estimation via normal equations
  - For linear problems, we need not calculate SS(β); rather, direct solution for *b* is possible
  - Recognize that vector of residuals will be normal to vector of *x* values at the least squares estimate

$$\begin{aligned} \sum (y - \hat{y})x &= 0 \\ \sum (y - bx)x &= 0 \\ \sum xy &= \sum bx^2 \end{aligned}$$

$$\Rightarrow \quad b = \frac{\sum xy}{\sum x^2}$$

- Estimate of experimental error
  - Assuming model structure is adequate, estimate s^2 of σ^2 can be obtained:

$$s^2 = \frac{SS_R}{n-1}$$

Precision of Estimate: Variance in *b*

- We can calculate the variance in our estimate of the slope, *b*:

$$b = \frac{\sum xy}{\sum x^2} \quad \Rightarrow \quad \hat{V}(b) = \frac{s^2}{\sum x_i^2}$$

$$\text{s.e.}(b) = \sqrt{\hat{V}(b)}$$

$$b \pm \text{s.e.}(b)$$

- Why?

$$\begin{aligned} b &= \frac{x_1}{\sum x^2} \cdot y_1 + \frac{x_2}{\sum x^2} \cdot y_2 + \cdots \frac{x_n}{\sum x^2} \cdot y_n \\ &= a_1 y_1 + a_2 y_2 + \cdots + a_n y_n \end{aligned}$$

$$\begin{aligned} V(b) &= (a_1^2 + a_2^2 + \cdots + a_n^2)\sigma^2 \\ &= \left[\left(\frac{x_1}{\sum x^2}\right)^2 + \cdots + \left(\frac{x_n}{\sum x^2}\right)^2 \right] \sigma^2 \\ &= \frac{\sum x^2}{(\sum x^2)^2} \sigma^2 \\ &= \frac{\sigma^2}{\sum x^2} \end{aligned}$$

Confidence Interval for β

- Once we have the standard error in *b*, we can calculate confidence intervals to some desired $(1-\alpha)100\%$ level of confidence

$$\frac{b-\beta}{\text{s.e.}(b)} \sim t \qquad \Rightarrow \qquad \beta = b \pm t_{\alpha/2} \cdot \text{s.e.}(b)$$

- Analysis of variance
  - Test hypothesis: $H_0 : \beta = b = 0$
  - If confidence interval for β includes 0, then β not significant

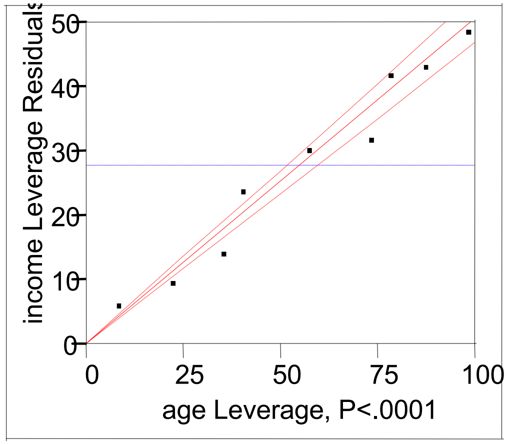
$$\begin{array}{ccccc} \sum y_i^2 & = & \sum \hat{y}_i^2 & + & \sum (y_i - \hat{y}_i)^2 \\ n & = & p & + & n-p \end{array}$$

  - Degrees of freedom (need in order to use t distribution)

p = # parameters estimated by least squares

Example Regression

Age	Income
8	6.16
22	9.88
35	14.35
40	24.06
57	30.34
73	32.17
78	42.18
87	43.23
98	48.76



Whole Model

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	8836.6440	8836.64	1093.146
Error	8	64.6695	8.08	Prob > F
C. Total	9	8901.3135		<.0001

Tested against reduced model: Y=0

Parameter Estimates

Term		Estimate	Std Error	t Ratio	Prob>|t|
Intercept	Zeroed	0	0	.	.
age		0.500983	0.015152	33.06	<.0001

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
age	1	1	8836.6440	1093.146	<.0001

- Note that this simple model assumes an intercept of zero – model must go through origin
- We can relax this requirement

Lack of Fit Error vs. Pure Error

- Sometimes we have replicated data
  - E.g. multiple runs at same *x* values in a designed experiment

- We can decompose the residual error contributions

$$SS_R = SS_L + SS_E$$

Where
SS_R = residual sum of squares error
SS_L = lack of fit squared error
SS_E = pure replicate error

- This allows us to TEST for lack of fit
  - By “lack of fit” we mean evidence that the linear model form is inadequate

$$\frac{s_L^2}{s_E^2} \sim F_{\nu_L,\nu_E}$$

Regression: Mean Centered Models

- Model form $y = \alpha + \beta(x - \bar{x})$
- Estimate by $\hat{y} = a + b(x - \bar{x}), \quad (y_i - \hat{y}_i) \sim \mathrm{N}(0, \sigma^2)$

$$\text{Minimize } SS_R = \sum (y_i - \hat{y}_i)^2 \text{ to estimate } \alpha \text{ and } \beta$$

$$a = \bar{y}$$

$$\mathrm{E}(a) = \alpha$$

$$\mathrm{Var}(a) = \mathrm{Var}\left[\frac{\sum y_i}{n}\right] = \frac{\sigma^2}{n}$$

$$b = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$\mathrm{E}(b) = \beta$$

$$\mathrm{Var}(b) = \frac{\sigma^2}{\sum (x_i - \bar{x})^2}$$

Regression: Mean Centered Models

- Confidence Intervals

$$\hat{y}_i = \bar{y} + b(x_i - \bar{x})$$

$$\begin{aligned} \mathrm{Var}(\hat{y}_i) &= \mathrm{Var}(\bar{y}) + (x_i - \bar{x})^2 \mathrm{Var}(b) \\ &= \frac{s^2}{n} + \frac{s^2 (x_i - \bar{x})^2}{\sum (x_i - \bar{x})^2} = s^2_{\hat{y}_i} \end{aligned}$$

- Our confidence interval on output *y* widens as we get further from the center of our data!

$$\hat{y}_i \pm t_{\alpha/2} \cdot s_{\hat{y}_i}$$

Polynomial Regression

- We may believe that a higher order model structure applies. Polynomial forms are also linear in the coefficients and can be fit with least squares

$$\eta = \beta_0 + \beta_1 x + \beta_2 x^2$$

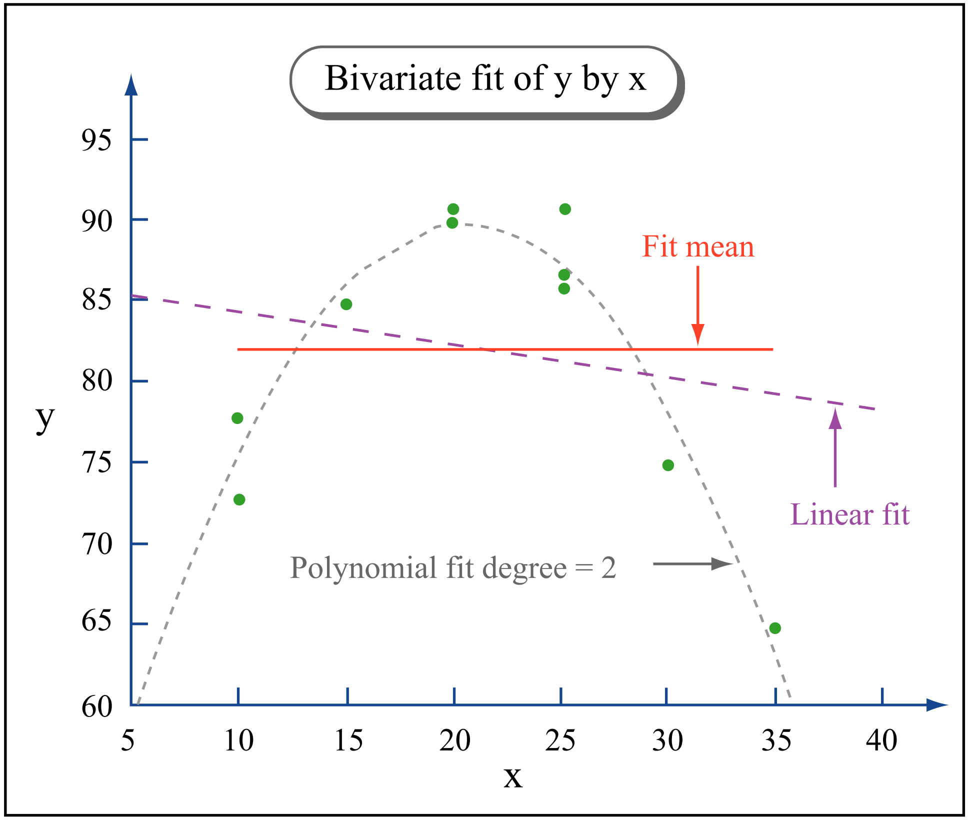
Curvature included through x^2 term

- Example: Growth rate data

Regression Example: Growth Rate Data

Observation Number	*Amount of Supplement (grams) x*	*Growth Rate (coded units) y*
1	10	73
2	10	78
3	15	85
4	20	90
5	20	91
6	25	87
7	25	86
8	25	91
9	30	75
10	35	65

Growth rate data



Figures by MIT OpenCourseWare.

- Replicate data provides opportunity to check for lack of fit

Growth Rate – First Order Model

- Mean significant, but linear term not
- Clear evidence of lack of fit

Source	*Sum of Squares*	*Degrees of Freedom*	*Mean Square*
Model	$S_M = 67{,}428.6$ { mean: 67,404.1; extra for linear: 24.5 }	2 { 1; 1 }	67,404.1 ← 24.5
→ Residual { lack of fit; pure error }	$S_R = 686.4$ { $S_L = 659.40$; $S_E = 27.0$ }	8 { 4; 4 }	85.8 { 164.85; 6.75 } ratio = 24.42
Total	$S_T = 68{,}115.0$	10	

Analysis of variance for growth rate data: Straight line model

Figure by MIT OpenCourseWare.

Growth Rate – Second Order Model

- No evidence of lack of fit
- Quadratic term significant

Source	*Sum of Squares*		*Degrees of Freedom*		*Mean Square*
Model	$S_M = 68{,}071.8$	mean 67,404.1	3	1	67,404.1
		extra for linear 24.5		1	24.5
		extra for quadratic 643.2		1	643.2
Residual	$S_R = 43.2$	$S_L = 16.2$	7	3	5.40 ratio = 0.80
		$S_E = 27.0$		4	6.75
Total	$S_T = 68{,}115.0$		10		

Analysis of variance for growth rate data: Quadratic model

Figure by MIT OpenCourseWare.

Polynomial Regression In Excel

- Create additional input columns for each input
- Use "Data Analysis" and "Regression" tool

x	x^2	y
10	100	73
10	100	78
15	225	85
20	400	90
20	400	91
25	625	87
25	625	86
25	625	91
30	900	75
35	1225	65

Regression Statistics	
Multiple R	0.968
R Square	0.936
Adjusted R Square	0.918
Standard Error	2.541
Observations	10

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	665.706	332.853	51.555	6.48E-05
Residual	7	45.194	6.456		
Total	9	710.9			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	35.657	5.618	6.347	0.0004	22.373	48.942
x	5.263	0.558	9.431	3.1E-05	3.943	6.582
x^2	-0.128	0.013	-9.966	2.2E-05	-0.158	-0.097

Polynomial Regression

- Generated using JMP package

Analysis of Variance

Source	DF	Sum of Square	Mean Squar	F Ratio
Model	2	665.70617	332.853	51.5551
Error	7	45.19383	6.456	Prob > F
C. Total	9	710.90000		<.0001

Lack Of Fit

Source	DF	Sum of Square	Mean Squar	F Ratio
Lack Of Fit	3	18.193829	6.0646	0.8985
Pure Error	4	27.000000	6.7500	Prob > F
Total Error	7	45.193829		0.5157
				Max RSq
				0.9620

Summary of Fit

RSquare	0.936427
RSquare Adj	0.918264
Root Mean Sq Error	2.540917
Mean of Response	82.1
Observations (or Sum Wgts)	10

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	35.657437	5.617927	6.35	0.0004
x	5.2628956	0.558022	9.43	<.0001
x*x	-0.127674	0.012811	-9.97	<.0001

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
x	1	1	574.28553	88.9502	<.0001
x*x	1	1	641.20451	99.3151	<.0001

Outline

- Response Surface Modeling (RSM)
  - Regression analysis, confidence intervals
- Process Optimization using DOE and RSM
  - Off-line/iterative
  - On-live/evolutionary

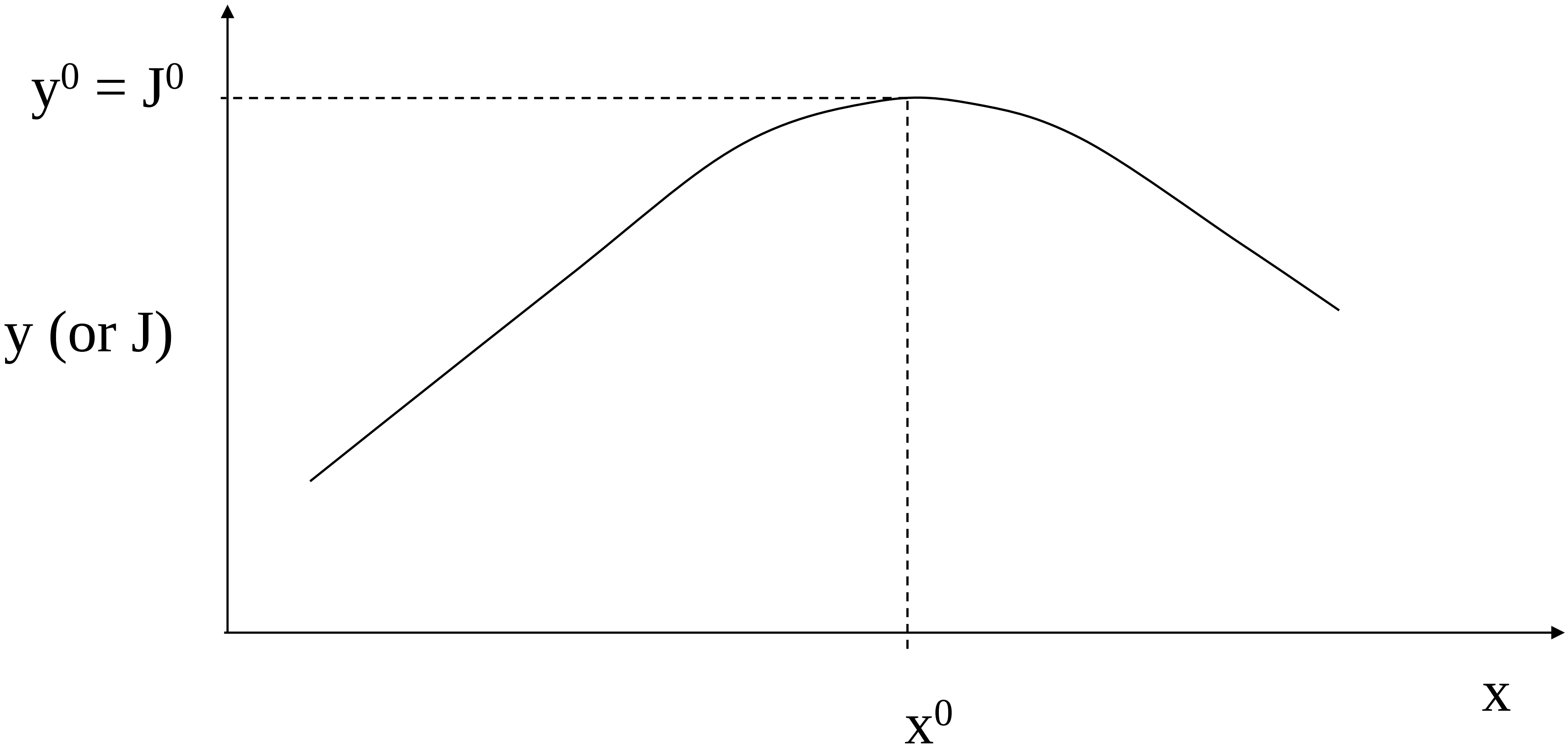
Process Optimization

- Multiple Goals in "Optimal" Process Output
  - Target mean for output(s) *Y*
  - Small variation/sensitivity $$\Delta Y = \frac{\partial Y}{\partial \alpha}\Delta\alpha + \frac{\partial Y}{\partial u}\Delta u$$
- Can Combine in an Objective Function "J"
  - Minimize or Maximize, e.g. $\min_{\underline{x}} J \quad \max_{\underline{x}} J$
  - Such that J = J(factors); might include J($\underline{x}$); J(α)
- Adjust J via factors with constraints

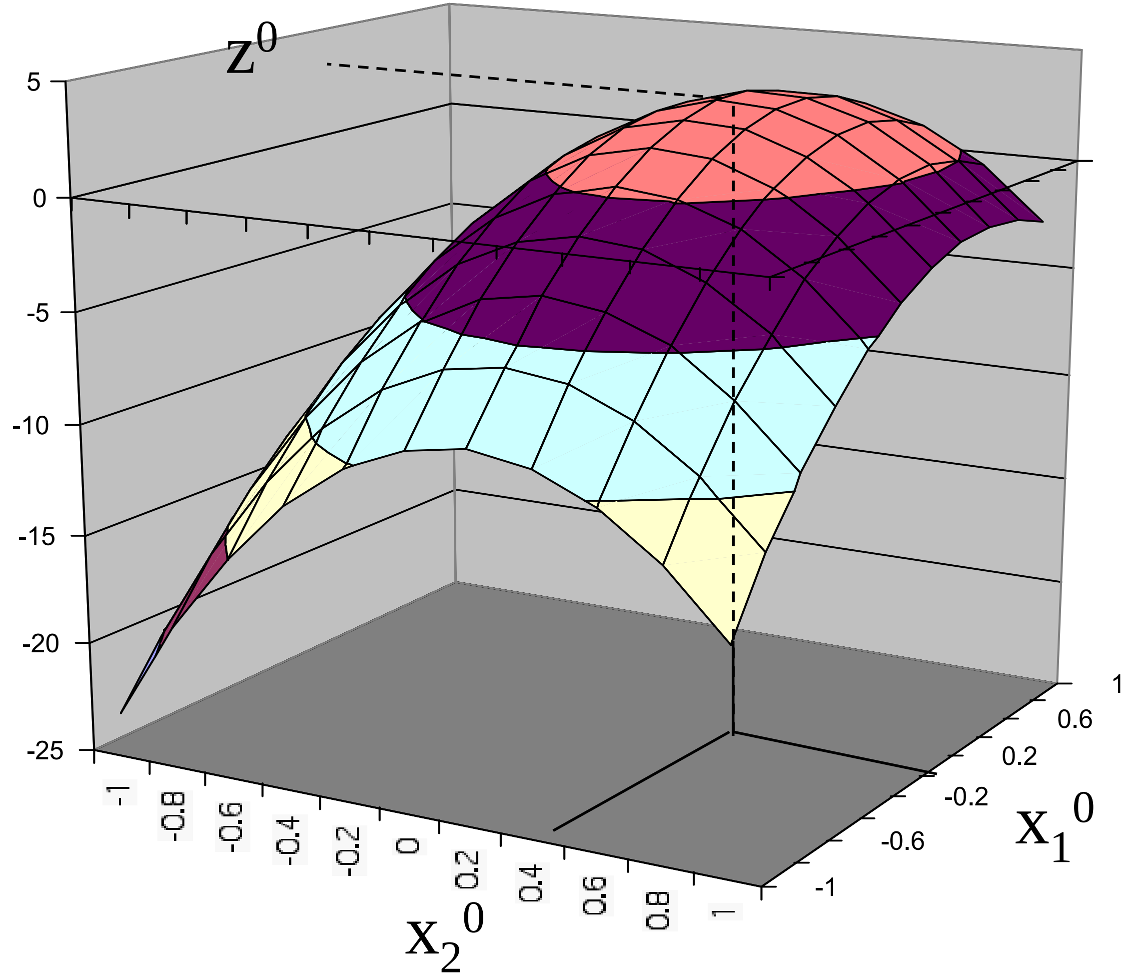
Methods for Optimization

- Analytical Solutions
  - $\partial y / \partial x = 0$
- Gradient Searches
  - Hill climbing (steepest ascent/descent)
  - Local min or max problem
  - Excel solver given a convex function
- Offline vs. Online

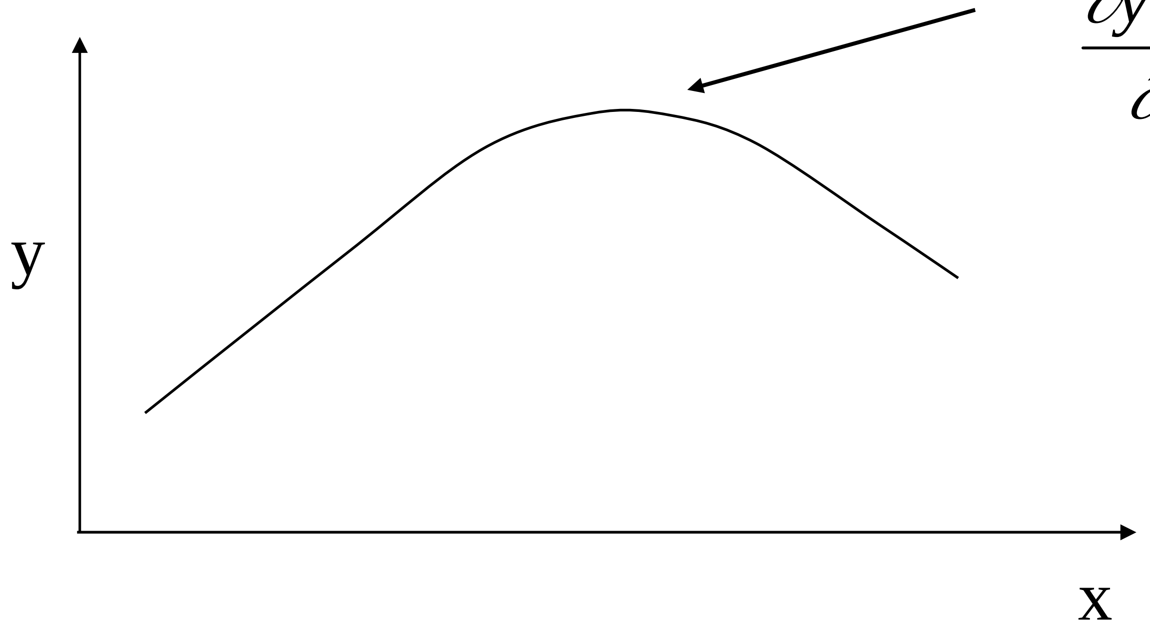
Basic Optimization Problem



3D Problem



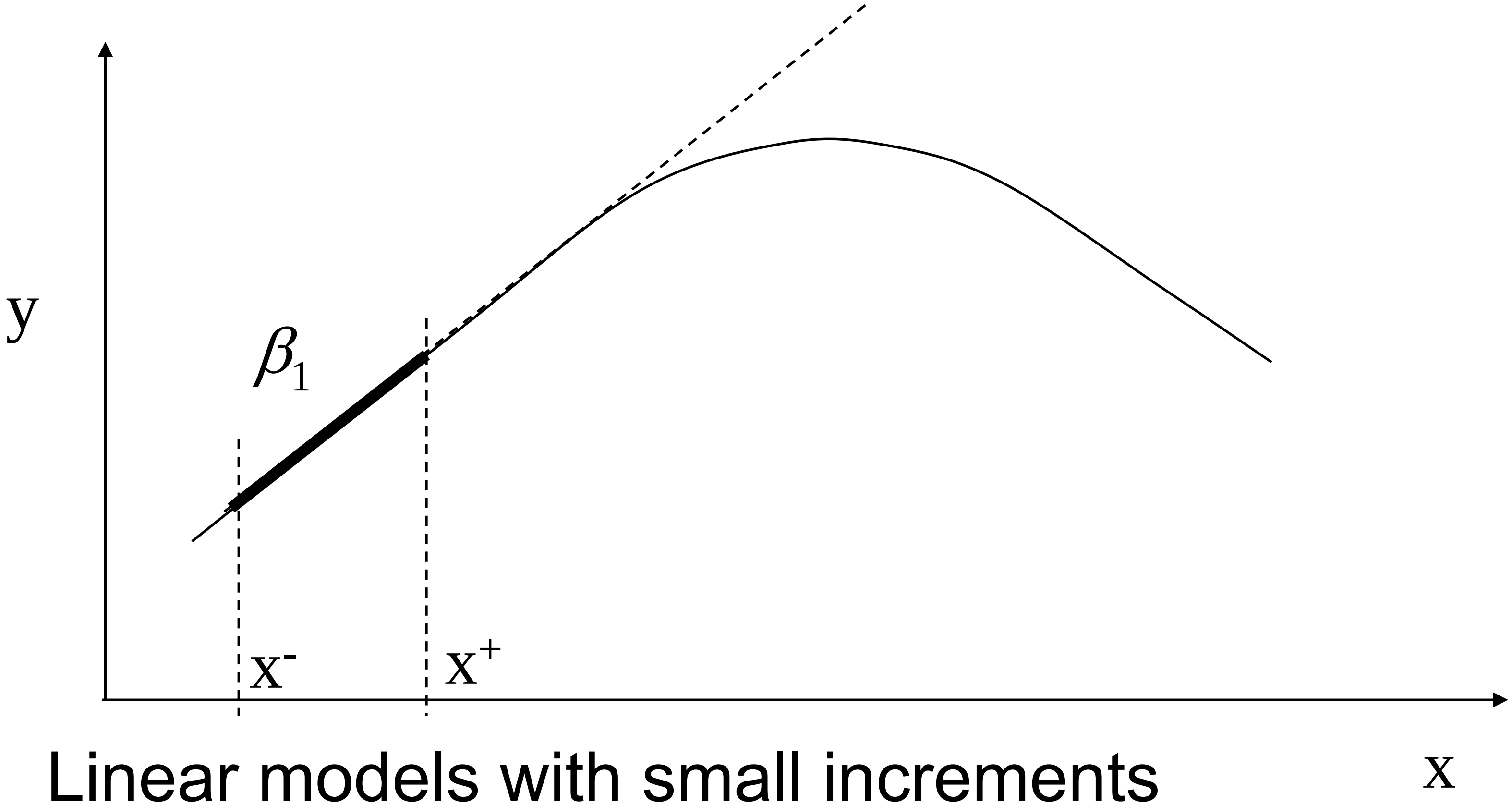
Analytical



$$\frac{\partial y(x)}{\partial x} = 0$$

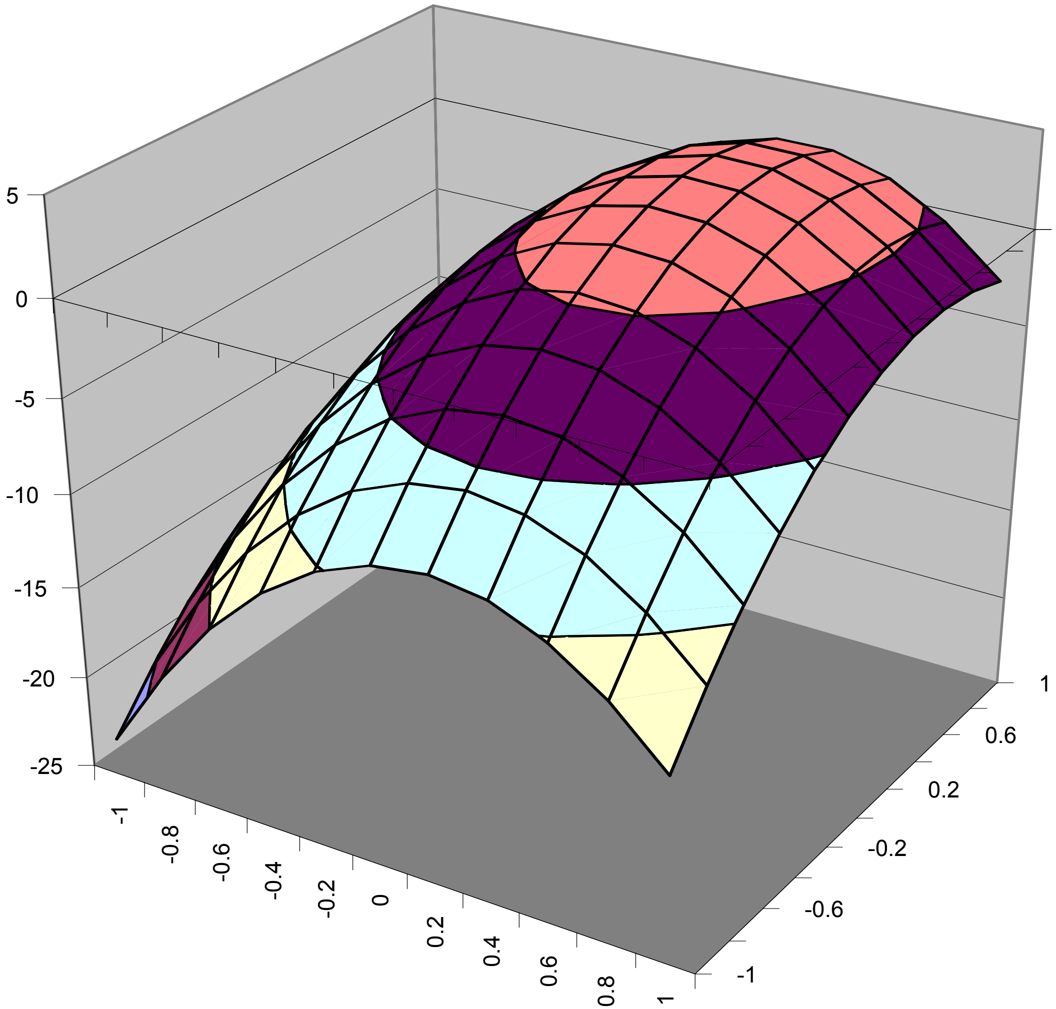
- Need Accurate *y(x)*
  - Analytical Model
  - Dense x increments in experiment
- Difficult with Sparse Experiments
  - Easy to missing optimum

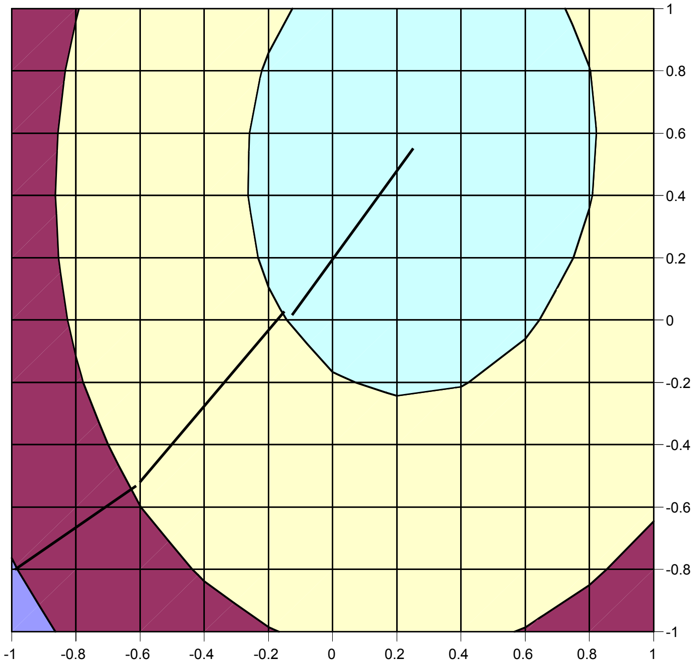
Sparse Data Procedure – Iterative Experiments/Model Construction



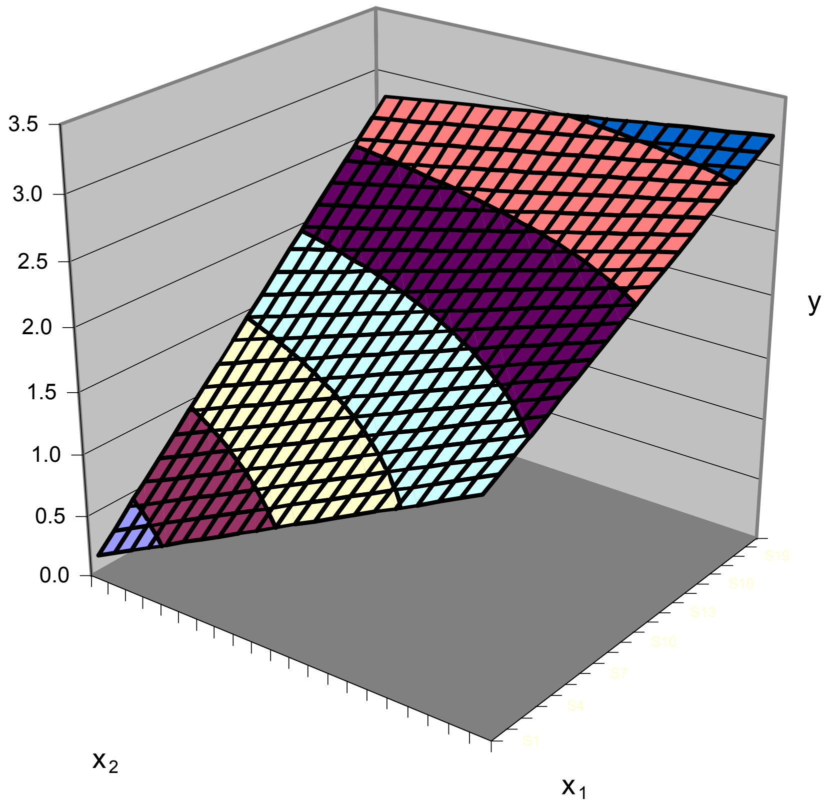
- Linear models with small increments
- Move along desired gradient
- Near zero slope change to quadratic model

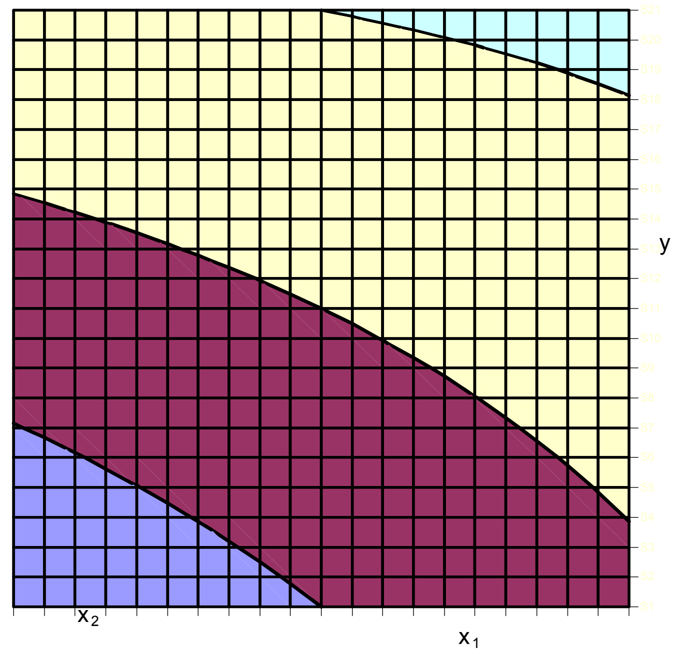
Extension to 3D





Linear Model Gradient Following





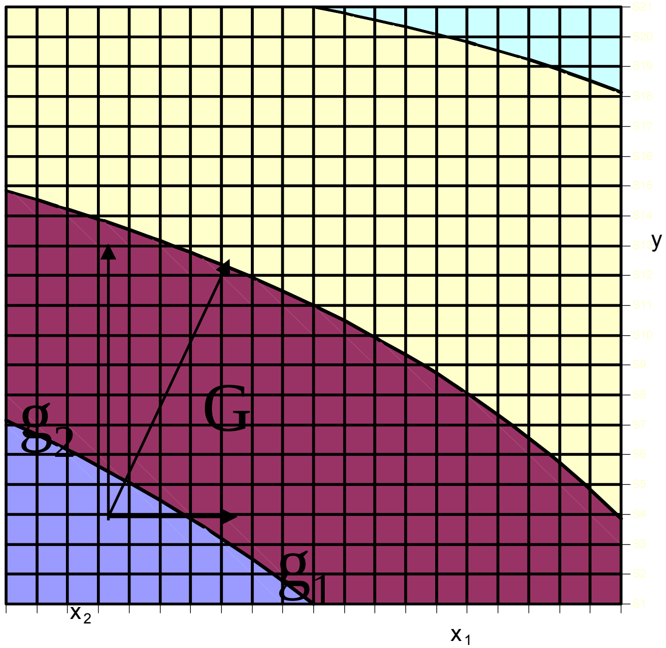
$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

Steepest Descent

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

$$g_{x_1} = \frac{\partial y}{\partial x_1} = \beta_1 + \beta_{12} x_2$$

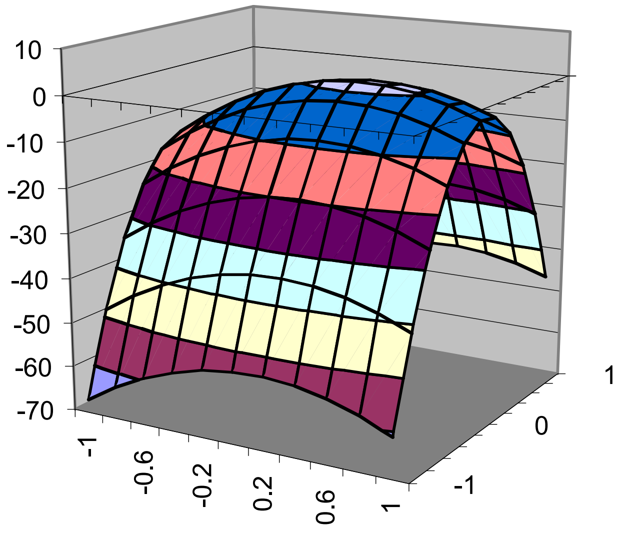
$$g_{x_2} = \frac{\partial y}{\partial x_2} = \beta_2 + \beta_{12} x_1$$

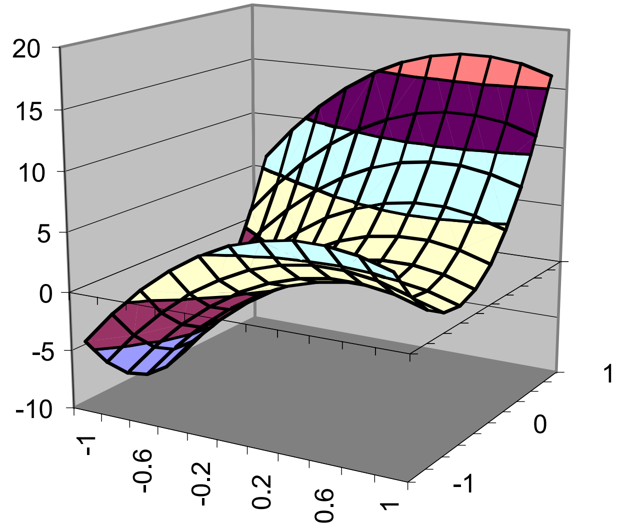


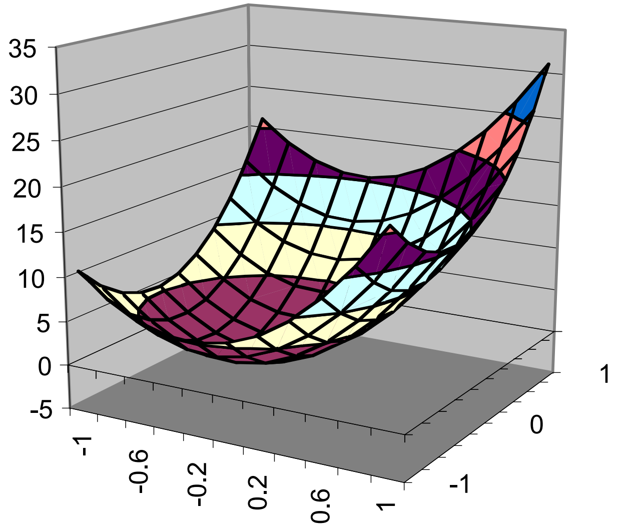
Make changes in x_1 and x_2 along G

$$\Delta x_2 = \frac{g_{x_1}}{g_{x_2}} \Delta x_1$$

Various Surfaces







A Procedure for DOE/Optimization

- Study Physics of Process
  - Define important inputs
  - Intuition about model
  - Limits on inputs
- DOE
  - Factor screening experiments
  - Further DOE as needed
  - RSM Construction
- Define Optimization/Penalty Function
  - J=f(x) $\max_{\underline{x}} J$ $\min_{\underline{x}} J$

For us, $\underline{x} = \underline{u}$ or $\underline{\alpha}$

(1) DOE Procedure

- Identify model (linear, quadratic, terms to include)
- Define inputs and ranges
- Identify "noise" parameters to vary if possible ($\Delta\alpha$'s)
- Perform experiment
  - – Appropriate order
    - randomization
    - blocking against nuisance or confounding effects

(2) RSM Procedure

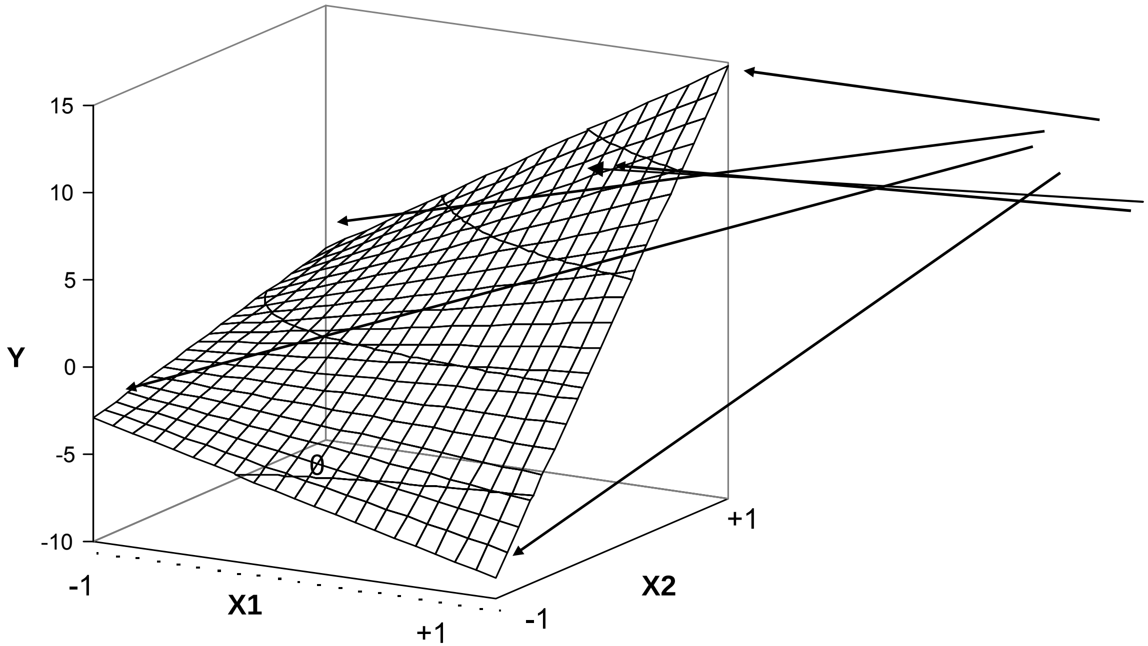
- Solve for $\underline{\beta}$'s
- Apply ANOVA
  - Data significant?
  - Terms significant?
  - Lack of Fit significant?
- Drop Insignificant Terms
- Add Higher Order Terms as needed

(3) Optimization Procedure

- Define Optimization/Penalty Function
- Search for Optimum
  - Analytically
  - Piecewise
  - Continuously/evolutionary
- Confirm Optimum

Confirming Experiments

- Checking intermediate points



- Data only at corners
- Test at interior point
- Evaluate error
- Consider Central Composite?

- Rechecking the “optimum”

Optimization Confirmation Procedure

- Find optimum value x*
- Perform confirming experiment
  - Test model at x*
  - Evaluate error with respect to model
  - Test hypothesis that $y(\underline{x}^*) = \hat{y}(\underline{x}^*)$
- If hypothesis fails
  - Consider new ranges for inputs
  - Consider higher order model as needed
  - Boundary may be optimum!

Experimental Optimization

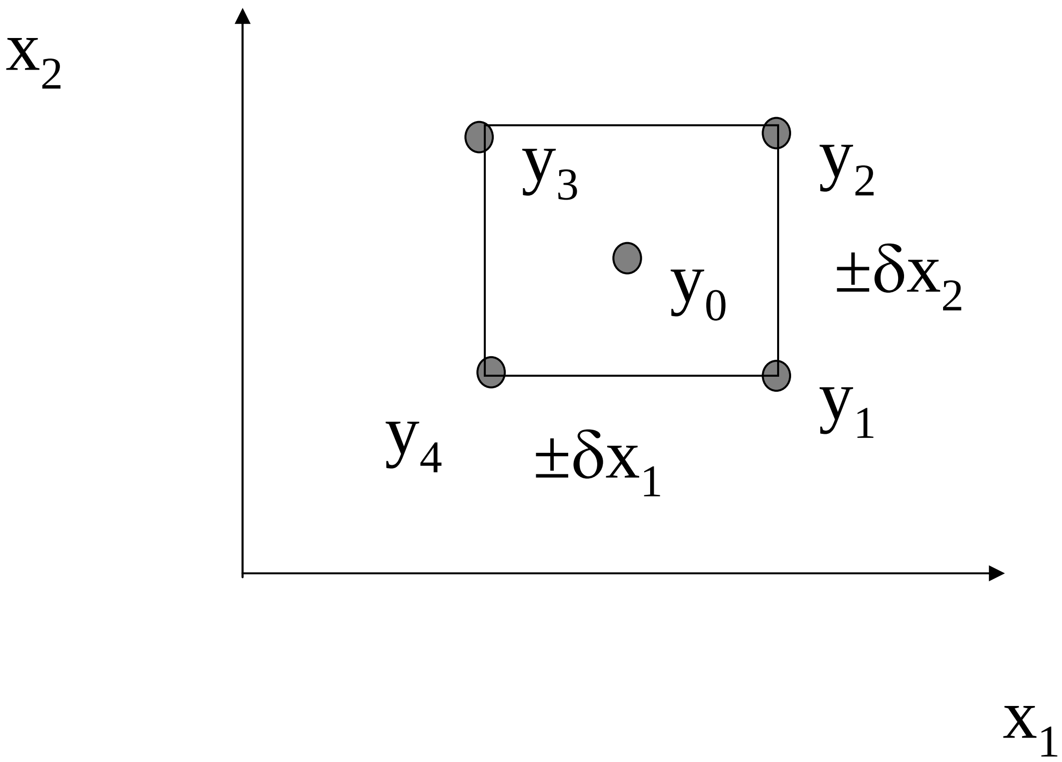
- WHY NOT JUST PICK BEST POINT?
- Why not optimize on-line?
  - Skip the Modeling Step?
- Adaptive Methods
  - Learn how best to model as you go
    - e.g. Adaptive OFACT

On-Line Optimization

- Perform 2^k Experiment
- Calculate Gradient
- Re-center 2^k Experiment About Maximum Corner
- Repeat
- Near Maximum?
  - – Should detect quadratic error
  - – Do quadratic fit near maximum point
    - Central Composite is good choice here
  - – Can also scale and rotate about principal axes

Continuous Optimization: EVOP

- Evolutionary Operation



- Pick "best" y_i
- Re-center process
- Do again

Summary

- Response Surface Modeling (RSM)
  - Regression analysis, confidence intervals
- Process Optimization using DOE and RSM
  - Off-line/iterative
  - On-live/evolutionary
- Next Time:
  - Process Robustness
  - Variation Modeling
  - Taguchi Approach