18-660: Numerical Methods for Engineering Design and Optimization

Xin Li
Department of ECE
Carnegie Mellon University
Pittsburgh, PA 15213

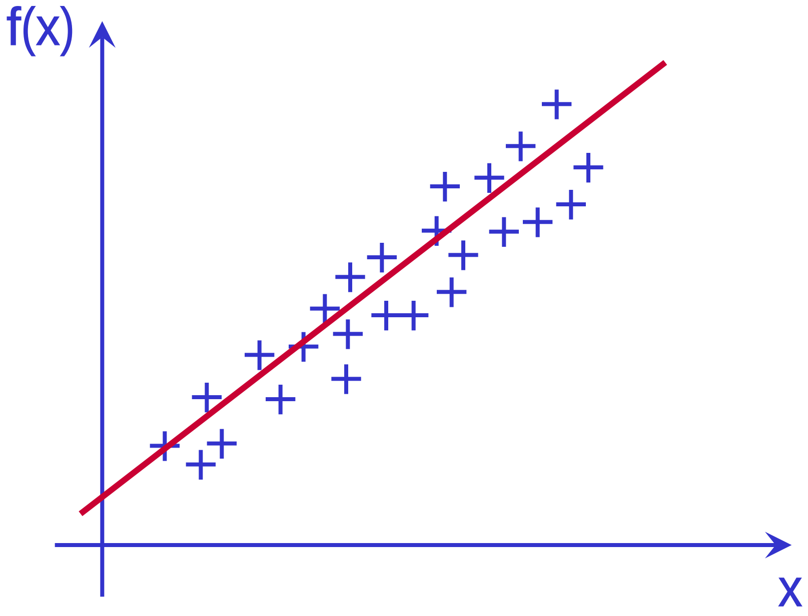
Electrical & Computer
ENGINEERING

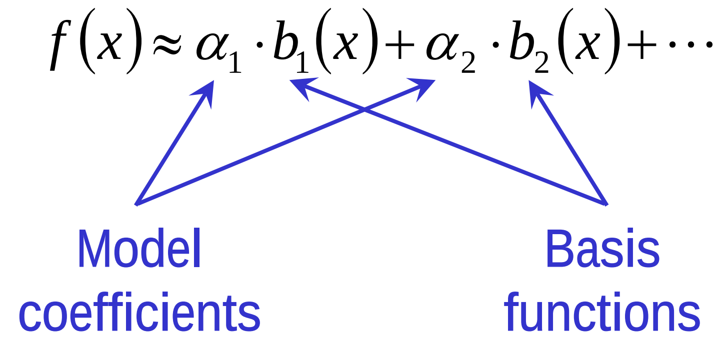
Overview

- Linear Regression
  - Ordinary least-squares regression
  - Minimax optimization
  - Design of experiments

Linear Regression

- Linear regression (also referred to as response surface modeling) is widely used for many engineering problems
  - We do not know the analytical form of f(x)
  - But we can generate a set of sampling points for f(x)
  - Fit an approximate function for f(x) from these sampling points





$$f(x) \approx \alpha_1 \cdot b_1(x) + \alpha_2 \cdot b_2(x) + \cdots$$

Model coefficients

Basis functions

f(x) is approximated as the linear combination of multiple basis functions

Linear Regression

- Major steps of linear regression
  - Select a model template (e.g., polynomial function)
  - Generate a number of sampling points
  - Compute performance values at these sampling points
  - Create a set of linear equations to solve model coefficients

- A simple example
  - $f(x) = \exp(x)$, $x \in [-1, 1]$
  - We will use this simple example to show how we can generally build a regression model from sampling data

Linear Regression Example

- Step 1: select a model template

$$f(x) \approx bx + c$$

- Step 2: generate a number of sampling points

Samples	1	2	3	4	5
x	-1	-0.5	0	0.5	1

- Step 3: compute performance values at these sampling points

Samples	1	2	3	4	5
f(x)	0.3679	0.6065	1.0000	1.6487	2.7183

Linear Regression Example

- Step 4: create linear equations for model coefficients

$$f(x) \approx bx + c$$

Samples	1	2	3	4	5
x	-1	-0.5	0	0.5	1
f(x)	0.3679	0.6065	1.0000	1.6487	2.7183

$$\begin{bmatrix} -1 & 1 \\ -0.5 & 1 \\ 0 & 1 \\ 0.5 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} b \\ c \end{bmatrix} = \begin{bmatrix} 0.3679 \\ 0.6065 \\ 1.0000 \\ 1.6487 \\ 2.7183 \end{bmatrix} \leftarrow \text{i-th sampling point}$$

x values

f(x) values

Linear Regression Example

- Step 5: solve over-determined linear equations
  - # of equations is greater than # of coefficients – over-determined
  - No exact solution exists to satisfy all equations, but we can find the least-squares solution:

$$A\cdot\alpha = B$$

$$\Downarrow$$

$$\min_{\alpha} \quad \underbrace{\left\|A\cdot\alpha - B\right\|_2^2}_{\text{Vector}}$$

Ordinary least-squares (OLS) regression

  - For a vector $\varepsilon \in R^M$, $||\varepsilon||_2$ is defined as:

$$\left\|\varepsilon\right\|_2 = \sqrt{\sum_{i=1}^{M} \varepsilon_i^2}$$

Linear Regression Example

$$A \cdot \alpha = B$$

i-th row

$$\begin{bmatrix} \\ \\ A \\ \\ \\ \end{bmatrix} \cdot \alpha - \begin{bmatrix} \\ \\ B \\ \\ \\ \end{bmatrix} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_M \end{bmatrix}$$

Error at the i-th sampling point

$$\min_{\alpha} \quad \left\| A \cdot \alpha - B \right\|_2^2 \quad \Rightarrow \quad \min_{\alpha} \quad \sum_{i=1}^{M} \varepsilon_i^2(\alpha)$$

Linear Regression Example

$$\text{M samples}\left\{\underbrace{\begin{bmatrix} \\ \quad A \quad \\ \\ \end{bmatrix}}_{\text{N coefficients}}\right. \cdot \alpha = \begin{bmatrix} \\ B \\ \\ \end{bmatrix} \qquad (M > N)$$

- There are several possible ways to solve over-determined linear equations for linear regression
  - We will explain these algorithms in detail in future lectures
  - For now, you can simply use “α = A\B” in MATLAB

Linear Regression Example

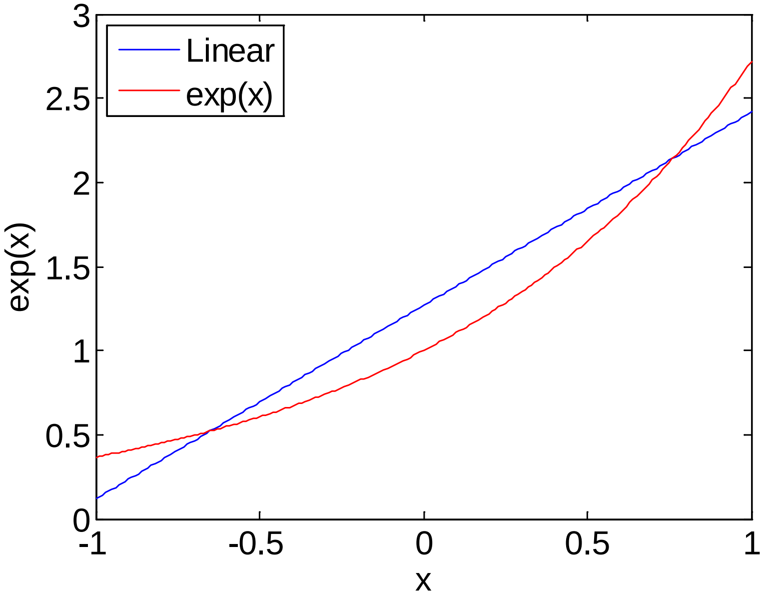
- Step 5: solve over-determined linear equations

$$\begin{bmatrix} -1 & 1 \\ -0.5 & 1 \\ 0 & 1 \\ 0.5 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} b \\ c \end{bmatrix} = \begin{bmatrix} 0.3679 \\ 0.6065 \\ 1.0000 \\ 1.6487 \\ 2.7183 \end{bmatrix}$$

$$\Downarrow$$

$$b = 1.1486$$

$$c = 1.2683$$



Linear model results in large error

Quadratic Model Example

- What if we build a quadratic model for y = exp(x)?
  - Select a model template

$$f(x) \approx ax^2 + bx + c$$

  - Generate a number of sampling points

Samples	1	2	3	4	5
x	-1	-0.5	0	0.5	1

  - Compute performance values at these sampling points

Samples	1	2	3	4	5
f(x)	0.3679	0.6065	1.0000	1.6487	2.7183

Quadratic Model Example

- Create a set of linear equations to solve model coefficients

$$f(x) \approx ax^2 + bx + c$$

Samples	1	2	3	4	5
x	-1	-0.5	0	0.5	1
f(x)	0.3679	0.6065	1.0000	1.6487	2.7183

$$\begin{array}{ccccc} \left[\begin{array}{ccc} 1 & -1 & 1 \\ 0.25 & -0.5 & 1 \\ 0 & 0 & 1 \\ 0.25 & 0.5 & 1 \\ 1 & 1 & 1 \end{array}\right] & \cdot & \begin{bmatrix} a \\ b \\ c \end{bmatrix} & = & \begin{bmatrix} 0.3679 \\ 0.6065 \\ 1.0000 \\ 1.6487 \\ 2.7183 \end{bmatrix} \\ \begin{array}{ccc} \downarrow & \downarrow & \\ x^2 & x & \end{array} & & & & \begin{array}{c} \downarrow \\ \text{f(x) values} \end{array} \end{array}$$

Quadratic Model Example

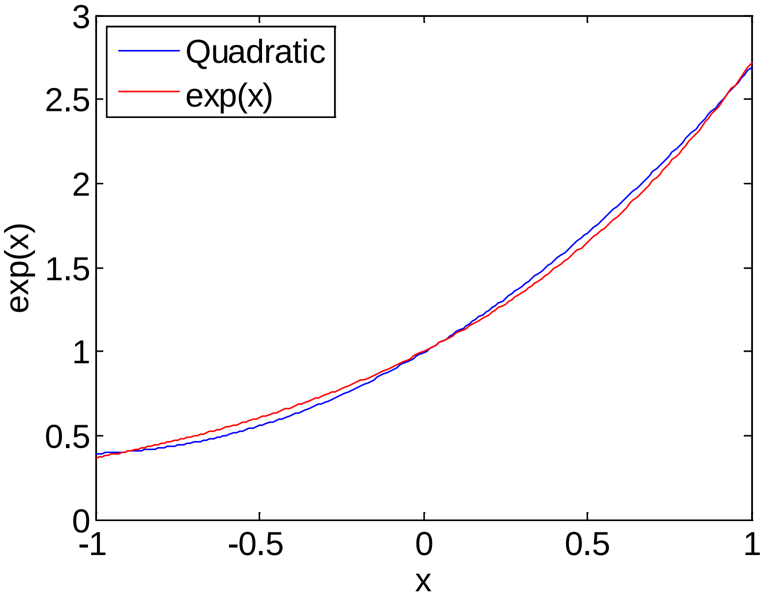
- Build quadratic model for y = exp(x)

$$\begin{bmatrix} 1 & -1 & 1 \\ 0.25 & -0.5 & 1 \\ 0 & 0 & 1 \\ 0.25 & 0.5 & 1 \\ 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0.3679 \\ 0.6065 \\ 1.0000 \\ 1.6487 \\ 2.7183 \end{bmatrix}$$

$$a = 0.5477$$

$$b = 1.1486$$

$$c = 0.9944$$



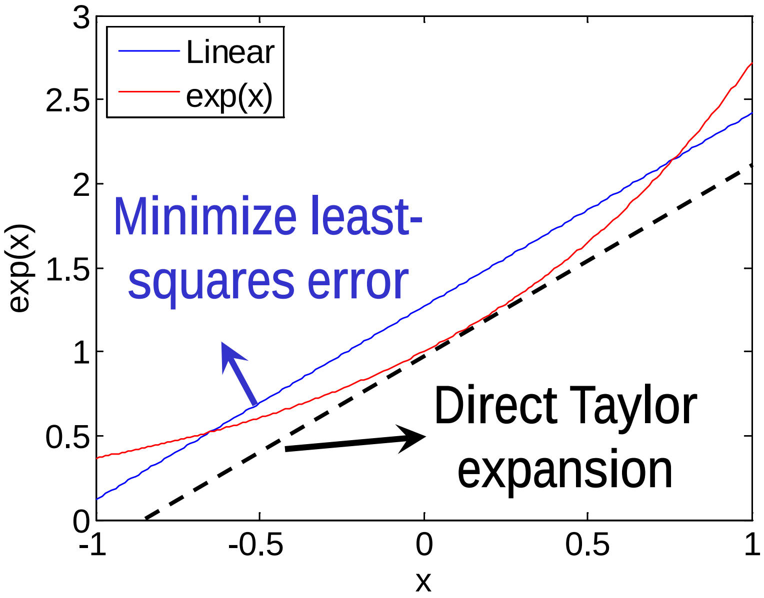
Quadratic model results in much better accuracy in this example

Linear Model vs. Quadratic Model

Linear RSM $\exp(x) \approx 1.1486x + 1.2683$

Quadratic RSM $\exp(x) \approx 0.5477x^2 + 1.1486x + 0.9944$

- Regression model is different from direct Taylor expansion
  - E.g., different constant terms in linear and quadratic models – they are selected to minimize the least-squares error



Linear model for exp(x)

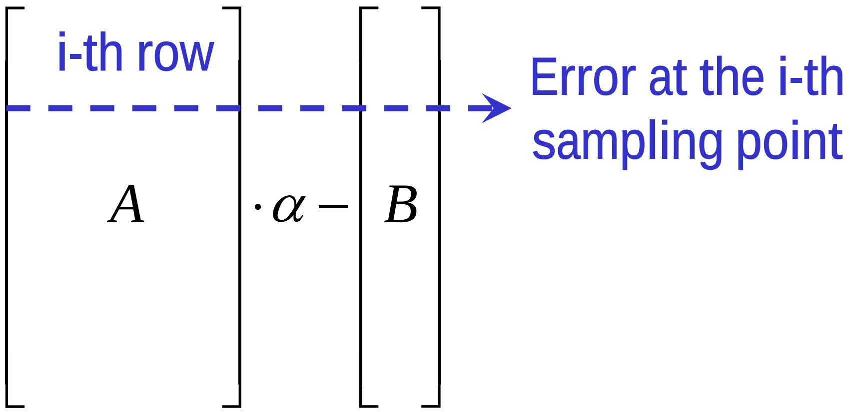
Minimax Optimization

- We can also solve over-determined linear equations to satisfy other optimality criteria (i.e., not ordinary least-squares)

$$A \cdot \alpha = B$$

$$\min_{\alpha} \quad \max_{i} \left| \overbrace{A(i,:)}^{\text{i-th row of A}} \cdot \alpha - B_i \right|$$

Minimize the maximal absolute error



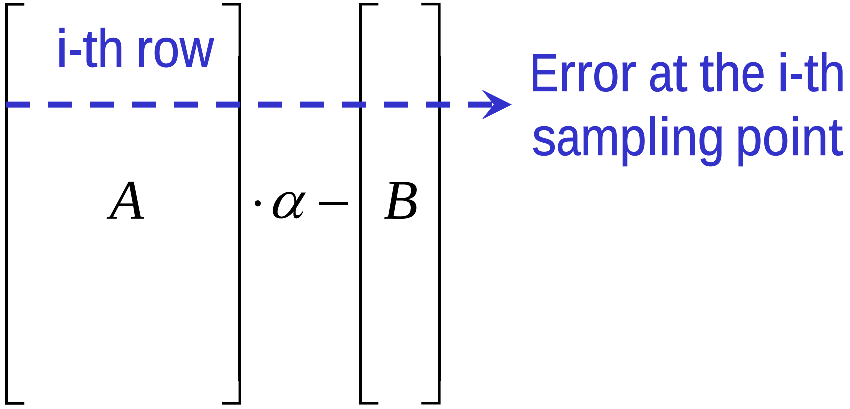
Minimax Optimization

- Other optimality criteria can be similarly formulated

$$A \cdot \alpha = B$$

$$\min_{\alpha} \quad \max_{i} \left| \frac{\overbrace{A(i,:)}^{\text{i-th row of A}} \cdot \alpha - B_i}{B_i} \right|$$

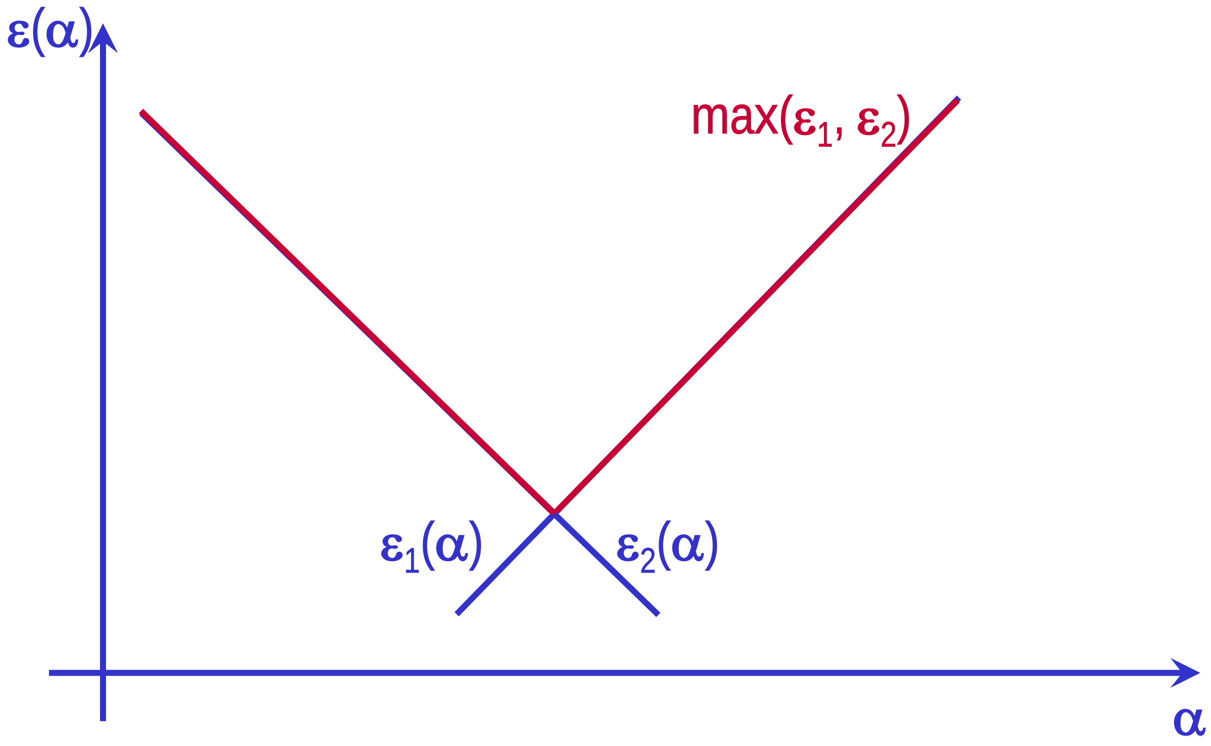
Minimize the maximal relative error



These formulations are minimax optimization problems

Minimax Optimization

- General minimax problems are difficult to solve
    - Cost function does not have continuous derivative



Minimax Optimization

- However, our minimax problem for regression modeling can be re-formulated into a special form

- Consider the example of absolute error minimization

$$\min_{\alpha} \quad \max_{i} \left| A(i,:) \cdot \alpha - B_i \right|$$

$$\Downarrow$$

Introduce a slack variable t

$$\min_{\alpha,t} \quad t$$

Cost function

$$\text{S.T.} \quad \begin{cases} \left| A(1,:) \cdot \alpha - B_1 \right| \le t \\ \left| A(2,:) \cdot \alpha - B_2 \right| \le t \\ \vdots \\ \left| A(M,:) \cdot \alpha - B_M \right| \le t \end{cases}$$

Constraints

Subject to

Minimax Optimization

$$\min_{\alpha,t} \quad t \qquad \text{S.T.} \quad \begin{cases} \left|A(1,:)\cdot\alpha - B_1\right| \le t \\ \left|A(2,:)\cdot\alpha - B_2\right| \le t \\ \vdots \\ \left|A(M,:)\cdot\alpha - B_M\right| \le t \end{cases}$$

$$\Longrightarrow$$

$$\min_{\alpha,t} \quad t \qquad \text{S.T.} \quad \begin{cases} -t \le A(1,:)\cdot\alpha - B_1 \le t \\ -t \le A(2,:)\cdot\alpha - B_2 \le t \\ \vdots \\ -t \le A(M,:)\cdot\alpha - B_M \le t \end{cases}$$

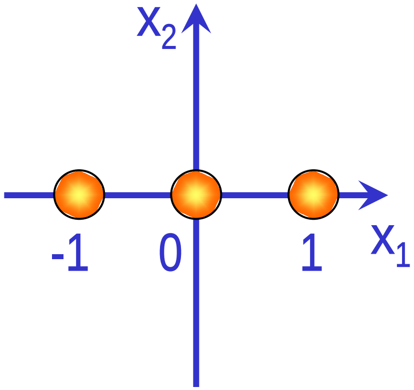
- Re-written as a linear programming (LP) problem
  - Both cost function and constraints are linear
  - No closed-form solution exists for LP
  - Can be numerically solved by an efficient (i.e., low complexity) and robust (i.e., global convergence) algorithm

Design of Experiments (DOE)

- We already know the basics for linear regression

- Open problem:
  - How can we select few samples to achieve good accuracy?

- A bad linear model example: $f(x_1, x_2) = a \cdot x_1 + b \cdot x_2 + c$



$$\begin{pmatrix} x_1 = -1 & x_2 = 0 & f_1 \end{pmatrix}$$
$$\begin{pmatrix} x_1 = 0 & x_2 = 0 & f_2 \end{pmatrix}$$
$$\begin{pmatrix} x_1 = 1 & x_2 = 0 & f_3 \end{pmatrix}$$

Sampling points for linear model

Design of Experiments (DOE)

- Linear model example (continued)

$$f\left(x_1, x_2\right) = a \cdot x_1 + b \cdot x_2 + c$$

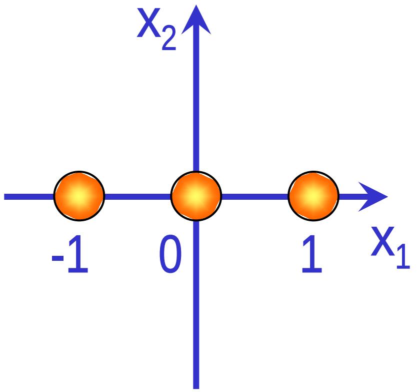
$$\begin{array}{l} \left(x_1 = -1 \quad x_2 = 0 \quad f_1\right) \\ \left(x_1 = 0 \quad x_2 = 0 \quad f_2\right) \\ \left(x_1 = 1 \quad x_2 = 0 \quad f_3\right) \end{array}$$

$$\begin{array}{c} \begin{array}{ccc} x_1 & x_2 & 1 \end{array} \\ \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \end{array} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}$$

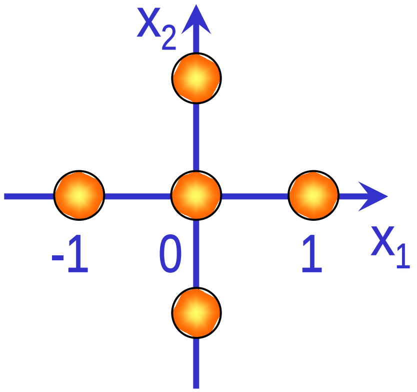
Singular matrix (cannot solve the coefficient b)

Design of Experiments (DOE)

- Linear model example (continued)



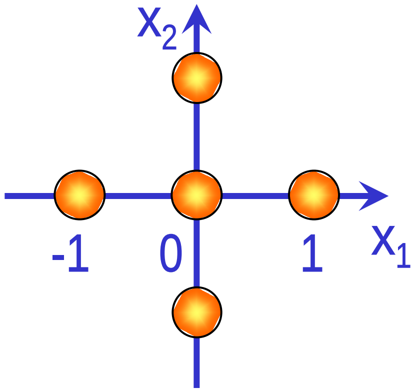
No variation is applied to x_2



Add additional sampling points for x_2

Design of Experiments (DOE)

- A bad quadratic model example: $f(x_1, x_2) = a_{11} \cdot x_1^2 + a_{12} \cdot x_1 \cdot x_2 + a_{22} \cdot x_2^2 + b_1 \cdot x_1 + b_2 \cdot x_2 + c$



$$\begin{pmatrix} x_1 = 0 & x_2 = 0 & f_1 \end{pmatrix}$$
$$\begin{pmatrix} x_1 = 0 & x_2 = -1 & f_2 \end{pmatrix}$$
$$\begin{pmatrix} x_1 = 0 & x_2 = 1 & f_3 \end{pmatrix}$$
$$\begin{pmatrix} x_1 = -1 & x_2 = 0 & f_4 \end{pmatrix}$$
$$\begin{pmatrix} x_1 = 1 & x_2 = 0 & f_5 \end{pmatrix}$$

Sampling points for quadratic model

Design of Experiments (DOE)

- Quadratic model example (continued)

$$f\left(x_1, x_2\right) = a_{11} \cdot x_1^2 + a_{12} \cdot x_1 \cdot x_2 + a_{22} \cdot x_2^2 \\ + b_1 \cdot x_1 + b_2 \cdot x_2 + c$$

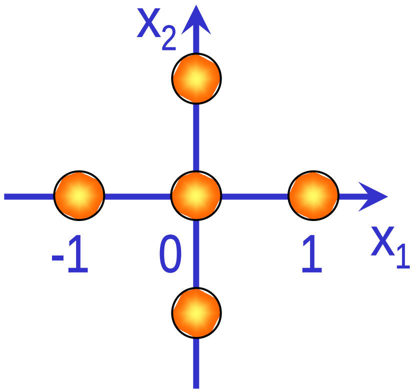
$$\left(x_1 = 0 \quad x_2 = 0 \quad f_1\right)$$
$$\left(x_1 = 0 \quad x_2 = -1 \quad f_2\right)$$
$$\left(x_1 = 0 \quad x_2 = 1 \quad f_3\right)$$
$$\left(x_1 = -1 \quad x_2 = 0 \quad f_4\right)$$
$$\left(x_1 = 1 \quad x_2 = 0 \quad f_5\right)$$

$$\begin{array}{c} \begin{array}{cccccc} x_1^2 & x_1x_2 & x_2^2 & x_1 & x_2 & 1 \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & -1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \end{array} \cdot \begin{bmatrix} a_{11} \\ a_{12} \\ a_{22} \\ b_1 \\ b_2 \\ c \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \end{bmatrix}$$

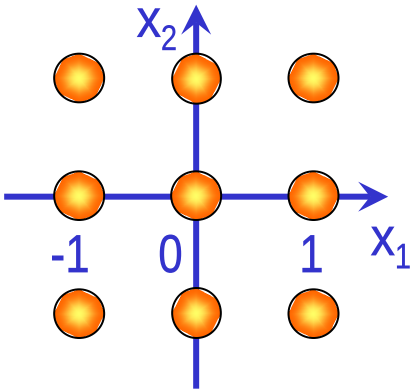
Singular matrix (cannot solve the coefficient a_{12})

Design of Experiments (DOE)

- Quadratic model example (continued)



Cross-product terms cannot be captured



Add additional sampling points for x_1x_2

Design of Experiments (DOE)

- Design of experiments (DOE) is a research area that studies how to optimally select sampling points for modeling

- Given a model template (e.g., linear or quadratic function), optimize sampling points for certain optimal criterion
  - E.g., maximize modeling accuracy

- Numerical optimization may be required to find the optimal sampling scheme

D. Montgomery, Design and Analysis of Experiments, John Wiley & Sons, 2004

Summary

- Linear regression
  - Ordinary least-squares regression
  - Minimax optimization
  - Design of experiments