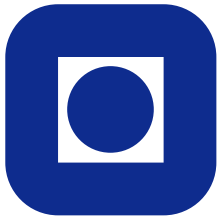

NTNU
Norwegian University of
Science and Technology

TMA4115 - Calculus 3

Week 17

Louise Meier Carlsen
Norwegian University of Science and Technology

Spring 2013

Review of last week

- Inner/dot product, norm/length, and orthogonality (Section 6.1),
- Orthogonal sets (Section 6.2),
- Orthogonal projections (Section 6.3)
- The Gram-Schmidt process (Section 6.4)

NTNU
Norwegian University of
Science and Technology

Plan for this week

This week we shall look at *least-squares problems*, applications to linear models, *diagonalization of symmetric matrices* and *quadratic forms*.

- Least-squares problems (Section 6.5),
- Applications to linear models (Section 6.6),
- Diagonalization of symmetric matrices (Section 7.1),
- Quadratic forms (Section 7.2).

NTNU
Norwegian University of
Science and Technology

Least-squares problems

- Let A be an $m \times n$ matrix and $\mathbf{b}$ a vector in $\mathbb{R}^m$.
- A *least-square solution* of $A\mathbf{x} = \mathbf{b}$ is an $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ such that

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$$

  for all $\mathbf{x}$ in $\mathbb{R}^n$.
- The number $\|\mathbf{b} - A\hat{\mathbf{x}}\|$ is called the *least-squares error* of $A\mathbf{x} = \mathbf{b}$.

NTNU
Norwegian University of
Science and Technology

Solution of the general least-squares problem

Theorem 13

The set of least-squares solutions of $A\mathbf{x} = \mathbf{b}$ coincides with the nonempty set of solutions of the normal equation $A^T A\mathbf{x} = A^T\mathbf{b}$.

NTNU
Norwegian University of
Science and Technology

Example

Let us find a least-squares solution of $A\mathbf{x} = \mathbf{b}$ where

$$A = \begin{bmatrix} 2 & 1 \\ -2 & 0 \\ 2 & 3 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} -5 \\ 8 \\ 1 \end{bmatrix}.$$

Example

Let us find all least-squares solutions of $A\mathbf{x} = \mathbf{b}$ where

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 7 \\ 2 \\ 3 \\ 6 \\ 5 \\ 4 \end{bmatrix}.$$

NTNU
Norwegian University of
Science and Technology

Uniqueness of least-squares solutions

Theorem 14

Let A be an $m \times n$ matrix. Then the following statements are logically equivalent.

1. The equation $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution for each $\mathbf{b}$ in $\mathbb{R}^m$.
2. The columns of A are linearly independent.
3. The matrix $A^T A$ is invertible.

When these statements are true, the least-squares solution $\hat{\mathbf{x}}$ is given by $\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b}$.

NTNU
Norwegian University of
Science and Technology

Example

Let us find the least-squares solution of $A\mathbf{x} = \mathbf{b}$ where $A = \begin{bmatrix} 1 & 2 \\ -1 & 4 \\ 1 & 2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}$.

QR-factorization and least-squares solutions

Theorem 15

Let A be an $m \times n$ matrix with linearly independent columns, let $A = QR$ be a QR-factorization of A and let $\mathbf{b}$ be in $\mathbb{R}^m$. Then $\hat{\mathbf{x}} = R^{-1}Q^T\mathbf{b}$ is the unique least-squares solution of $A\mathbf{x} = \mathbf{b}$.

NTNU
Norwegian University of
Science and Technology

Example

Suppose that $A = \begin{bmatrix} 1 & -1 \\ 1 & 4 \\ 1 & -1 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \\ 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & 5 \end{bmatrix}$.

Let us find the least-squares solution of $A\mathbf{x} = \mathbf{b}$ where

$\mathbf{b} = \begin{bmatrix} -1 \\ 6 \\ 5 \\ 7 \end{bmatrix}$.

NTNU
Norwegian University of
Science and Technology

Least-squares lines

- Suppose we are given a finite number of points $(x_1, y_1), \ldots, (x_n, y_n)$.
- We want to determine the line $y = \beta_0 + \beta_1 x$ which is as "close" to the points as possible.
- For each point (x_j, y_j) we call
  - y_j for the *observed* value of y,
  - $\beta_0 + \beta_1 x_j$ for the *predicted* value of y,
  - $y_j - (\beta_0 + \beta_1 x_j)$ for the *residual*.
- The *least-squares line* is the line $y = \beta_0 + \beta_1 x$ that minimizes the sum of the squares of the residuals.

NTNU
Norwegian University of
Science and Technology

Finding the least-squares line

Finding the least-squares line is equivalent to computing the least-squares solution of the equation $X\beta = \mathbf{y}$ where

$$X = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \text{ and } \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

NTNU
Norwegian University of
Science and Technology

Example

Let us determine the equation $y = \beta_0 + \beta_1 x$ of the least-squares line that best fits the data points (1,0), (2,1), (4,2) and (5,3).

The general linear model

- Any equation of the form $X\beta = \mathbf{y} + \epsilon$ where X and $\mathbf{y}$ are given, is referred to as a *linear model*.
- The goal is to minimize the length of the *residual vector* ϵ.
- This amounts to finding a least-squares solution of $X\beta = \mathbf{y}$.

NTNU
Norwegian University of
Science and Technology

Example

Suppose data points $(x_1, y_1), \ldots, (x_n, y_n)$ appear to lie along a parabola.
Suppose we wish to approximate the data by an equation of the form $y = \beta_0 + \beta_1 x + \beta_2 x^2$.
Let us describe the linear model that produces a least-square fit of the data point to the above equation.

Solution

Let $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$, $X = \begin{bmatrix} 1 & x_1 & x_1^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix}$, $\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$ and

$$\epsilon = \begin{bmatrix} y_1 - (\beta_0 + \beta_1 x_1 + \beta_2 x_1^2) \\ \vdots \\ y_n - (\beta_0 + \beta_1 x_n + \beta_2 x_n^2) \end{bmatrix}.$$

Then $X\beta = \mathbf{y} + \epsilon$ is the linear model that produces a least-square fit of the data point to the equation $y = \beta_0 + \beta_1 x + \beta_2 x^2$.

NTNU
Norwegian University of
Science and Technology

Example

Suppose we are given data points $(u_1, v_1, y_1), \ldots, (u_n, v_n, y_n)$ that we expect to satisfy an equation of the form $y = \beta_0 + \beta_1 u + \beta_2 v$.
Let us describe the linear model that produces a least-square fit of the data point to the above equation.

NTNU
Norwegian University of
Science and Technology

Solution

Let $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$, $X = \begin{bmatrix} 1 & u_1 & v_1 \\ \vdots & \vdots & \vdots \\ 1 & u_n & v_n \end{bmatrix}$, $\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$ and

$$\epsilon = \begin{bmatrix} y_1 - (\beta_0 + \beta_1 u_1 + \beta_2 v_1) \\ \vdots \\ y_n - (\beta_0 + \beta_1 u_n + \beta_2 v_n) \end{bmatrix}.$$

Then $X\beta = \mathbf{y} + \epsilon$ is the linear model that produces a least-square fit of the data point to the equation $y = \beta_0 + \beta_1 u + \beta_2 v$.

NTNU
Norwegian University of
Science and Technology