KUTAISI INTERNATIONAL UNIVERSITY

Numerical Linear Algebra

Ramaz Botchorishvili

Kutaisi International University

December 14, 2022

KUTAISI
INTERNATIONAL
UNIVERSITY

Numerical Linear Algebra

Ramaz Botchorishvili

Kutaisi International University

December 14, 2022

QR factorization, Least squares

- QR and reduced QR factorization
- Linear least squares
- Constrained least squares
- Q & A

Recap of Previous Lecture

- Recap of Previous Lecture
- Method of minimal residuals
- Steepest Descent
- Gram-Schmidt orthogonalization

QR factorization

Definition 13.1

QR Fcatorization of rectangular matrix:

- $A \in \mathcal{R}^{n \times m}$

QR factorization

Definition 13.1

QR Fcatorization of rectangular matrix:

- $A \in \mathcal{R}^{n \times m}$
- $\exists$ orthogonal $Q \in \mathcal{R}^{n \times n}, Q^T Q = I$

QR factorization

Definition 13.1

QR Fcatorization of rectangular matrix:

- $A \in \mathcal{R}^{n\times m}$
- $\exists$ orthogonal $Q \in \mathcal{R}^{n\times n}, Q^TQ = I$
- $\exists$ upper trapezoidal R :

QR factorization

Definition 13.1

QR Fcatorization of rectangular matrix:

- $A \in \mathcal{R}^{n \times m}$
- $\exists$ orthogonal $Q \in \mathcal{R}^{n \times n}, Q^T Q = I$
- $\exists$ upper trapezoidal R :
  - $R \in \mathcal{R}^{m \times m}$

QR factorization

Definition 13.1

QR Fcatorization of rectangular matrix:

- $A \in \mathcal{R}^{n \times m}$
- $\exists$ orthogonal $Q \in \mathcal{R}^{n \times n}, Q^T Q = I$
- $\exists$ upper trapezoidal R :
  - $R \in \mathcal{R}^{m \times m}$
  - $r_{ij} = 0, \text{ for } i \geq n + 1$

QR factorization

Definition 13.1

QR Fcatorization of rectangular matrix:

- $A \in \mathcal{R}^{n \times m}$
- $\exists$ orthogonal $Q \in \mathcal{R}^{n \times n}, Q^T Q = I$
- $\exists$ upper trapezoidal R :
  - $R \in \mathcal{R}^{m \times m}$
  - $r_{ij} = 0, \text{ for } i \geq n+1$
- $A = QR$

QR factorization

Definition 13.1

QR Fcatorization of rectangular matrix:

- $A \in \mathcal{R}^{n \times m}$
- $\exists$ orthogonal $Q \in \mathcal{R}^{n \times n}, Q^T Q = I$
- $\exists$ upper trapezoidal R :
  - $R \in \mathcal{R}^{m \times m}$
  - $r_{ij} = 0, \text{ for } i \geq n + 1$
- $A = QR$

Definition 13.2

Reduced QR fatorization of rectangular matrix:

QR factorization

Definition 13.1

QR Fcatorization of rectangular matrix:

- $A \in \mathcal{R}^{n \times m}$
- $\exists$ orthogonal $Q \in \mathcal{R}^{n \times n}, Q^T Q = I$
- $\exists$ upper trapezoidal R :
  - $R \in \mathcal{R}^{m \times m}$
  - $r_{ij} = 0, \text{ for } i \geq n + 1$
- $A = QR$

Definition 13.2

Reduced QR fatorization of rectangular matrix:

- $A \in \mathcal{R}^{n \times m}$

QR factorization

Definition 13.1

QR Fcatorization of rectangular matrix:

- $A \in \mathcal{R}^{n\times m}$
- $\exists$ orthogonal $Q \in \mathcal{R}^{n\times n}, Q^TQ = I$
- $\exists$ upper trapezoidal R :
  - $R \in \mathcal{R}^{m\times m}$
  - $r_{ij} = 0,$ for $i \geq n+1$
- $A = QR$

Definition 13.2

Reduced QR fatorization of rectangular matrix:

- $A \in \mathcal{R}^{n\times m}$
- $\exists$ orthogonal $Q \in \mathcal{R}^{n\times m}, Q^TQ = I$

QR factorization

Definition 13.1

QR Fcatorization of rectangular matrix:

- $A \in \mathcal{R}^{n \times m}$
- $\exists$ orthogonal $Q \in \mathcal{R}^{n \times n}, Q^T Q = I$
- $\exists$ upper trapezoidal R :
  - $R \in \mathcal{R}^{m \times m}$
  - $r_{ij} = 0, \text{ for } i \geq n + 1$
- $A = QR$

Definition 13.2

Reduced QR fatorization of rectangular matrix:

- $A \in \mathcal{R}^{n \times m}$
- $\exists$ orthogonal $Q \in \mathcal{R}^{n \times m}, Q^T Q = I$
- $\exists$ triangular $R \in \mathcal{R}^{m \times m}$

QR factorization

Definition 13.1

QR Fcatorization of rectangular matrix:

- $A \in \mathcal{R}^{n \times m}$
- $\exists$ orthogonal $Q \in \mathcal{R}^{n \times n}, Q^T Q = I$
- $\exists$ upper trapezoidal R :
  - $R \in \mathcal{R}^{m \times m}$
  - $r_{ij} = 0,$ for $i \geq n + 1$
- $A = QR$

Definition 13.2

Reduced QR fatorization of rectangular matrix:

- $A \in \mathcal{R}^{n \times m}$
- $\exists$ orthogonal $Q \in \mathcal{R}^{n \times m}, Q^T Q = I$
- $\exists$ triangular $R \in \mathcal{R}^{m \times m}$
- $A = QR$

QR factorization

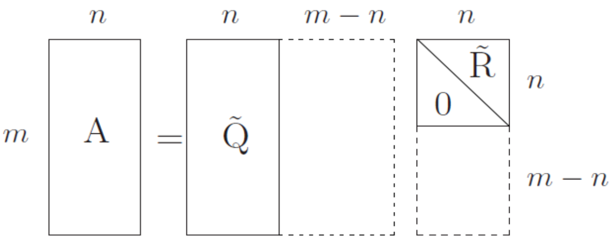


Figure: $A = QR$, $A = \tilde{Q}\tilde{R}$ - reduced QR factorization. From Quarteroni et al.

QR factorization

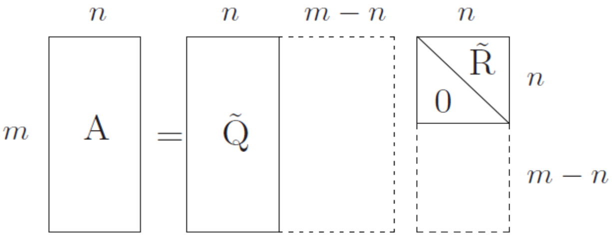


Figure: $A = QR$, $A = \tilde{Q}\tilde{R}$ - reduced QR factorization. From Quarteroni et al.

Theorem 13.3

- $A \in \mathcal{R}^{n\times m}$

QR factorization

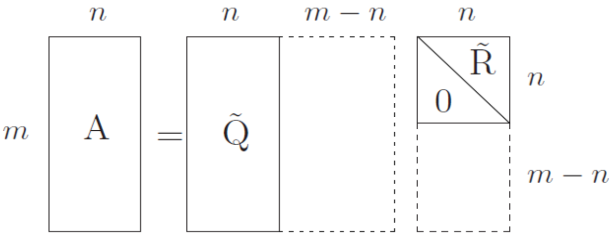


Figure: $A = QR$, $A = \tilde{Q}\tilde{R}$ - reduced QR factorization. From Quarteroni et al.

Theorem 13.3

- $A \in \mathcal{R}^{n \times m}$
- $\textbf{\textit{rank}}(A) = m$

QR factorization

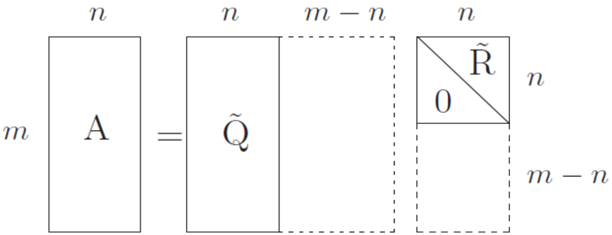


Figure: $A = QR$, $A = \tilde{Q}\tilde{R}$ - reduced QR factorization. From Quarteroni et al.

Theorem 13.3

- $A \in \mathcal{R}^{n\times m}$
- $\boldsymbol{rank}(A) = m$
- $\exists QR$ *factorization*

QR factorization

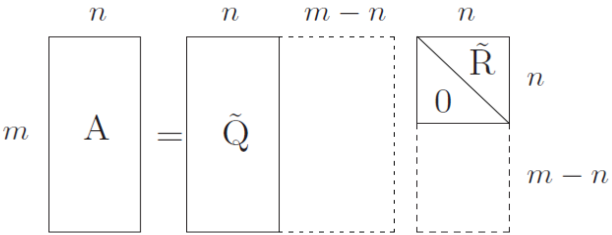


Figure: $A = QR$, $A = \tilde{Q}\tilde{R}$ - reduced QR factorization. From Quarteroni et al.

Theorem 13.3

- $A \in \mathcal{R}^{n \times m}$
- $\textbf{\textit{rank}}(A) = m$
- $\exists QR$ *factorization*
- $\Downarrow$

QR factorization

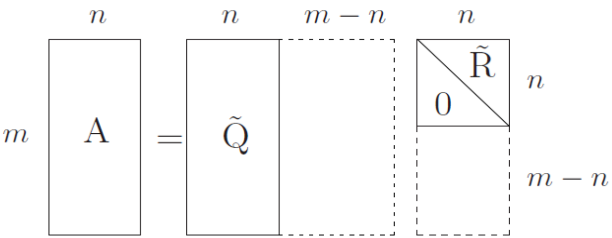
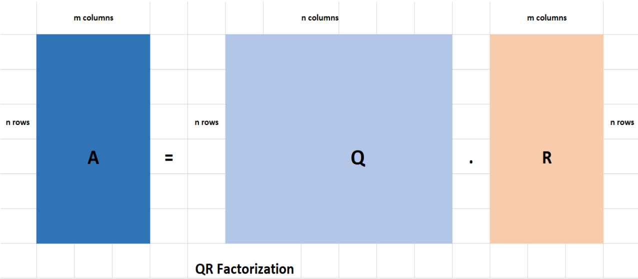


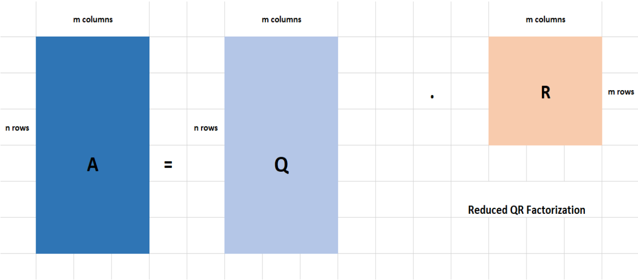
Figure: $A = QR$, $A = \tilde{Q}\tilde{R}$ - reduced QR factorization. From Quarteroni et al.

Theorem 13.3

- $A \in \mathcal{R}^{n \times m}$
- $\boldsymbol{rank}(A) = m$
- $\exists QR$ *factorization*
- $\Downarrow$
- $\exists$ *reduced* QR *factorization* $A = \tilde{Q}\tilde{R}$

QR and reduced QR





QR factorization

Example 13.4

```
A= [[ 1  2  3]
 [ 4  5  6]
 [ 7  8  9]
 [10 11 12]]
Q= [[-0.07761505 -0.83305216  0.53358462]
 [-0.31046021 -0.45123659 -0.8036038 ]
 [-0.54330537 -0.06942101  0.00645373]
 [-0.77615053  0.31239456  0.26356544]]
R= [[-1.28840987e+01 -1.45916299e+01 -1.62991610e+01]
 [ 0.00000000e+00 -1.04131520e+00 -2.08263040e+00]
 [ 0.00000000e+00  0.00000000e+00 -3.39618744e-15]]
```

Figure: Reduced QR factorization

QR factorization

Example 13.4

```
A= [[ 1  2  3]
 [ 4  5  6]
 [ 7  8  9]
 [10 11 12]]
Q= [[-0.07761505 -0.83305216  0.53358462]
 [-0.31046021 -0.45123659 -0.8036038 ]
 [-0.54330537 -0.06942101  0.00645373]
 [-0.77615053  0.31239456  0.26356544]]
R= [[-1.28840987e+01 -1.45916299e+01 -1.62991610e+01]
 [ 0.00000000e+00 -1.04131520e+00 -2.08263040e+00]
 [ 0.00000000e+00  0.00000000e+00 -3.39618744e-15]]
```

Figure: Reduced QR factorization

- A - tall matrix

QR factorization

Example 13.4

```
A= [[ 1  2  3]
 [ 4  5  6]
 [ 7  8  9]
 [10 11 12]]
Q= [[-0.07761505 -0.83305216  0.53358462]
 [-0.31046021 -0.45123659 -0.8036038 ]
 [-0.54330537 -0.06942101  0.00645373]
 [-0.77615053  0.31239456  0.26356544]]
R= [[-1.28840987e+01 -1.45916299e+01 -1.62991610e+01]
 [ 0.00000000e+00 -1.04131520e+00 -2.08263040e+00]
 [ 0.00000000e+00  0.00000000e+00 -3.39618744e-15]]
```

Figure: Reduced QR factorization

- A - tall matrix
- Q - tall matrix, orthogonal

QR factorization

Example 13.4

```
A= [[ 1  2  3]
 [ 4  5  6]
 [ 7  8  9]
 [10 11 12]]
Q= [[-0.07761505 -0.83305216  0.53358462]
 [-0.31046021 -0.45123659 -0.8036038 ]
 [-0.54330537 -0.06942101  0.00645373]
 [-0.77615053  0.31239456  0.26356544]]
R= [[-1.28840987e+01 -1.45916299e+01 -1.62991610e+01]
 [ 0.00000000e+00 -1.04131520e+00 -2.08263040e+00]
 [ 0.00000000e+00  0.00000000e+00 -3.39618744e-15]]
```

Figure: Reduced QR factorization

- A - tall matrix
- Q - tall matrix, orthogonal
- R - square matrix, upper triangular

Reduced QR factorization, 1

- Input: $A = (a_{ij})_{n \times k}$,
- A is tall matrix, $n > k$, full rank

Reduced QR factorization, 1

- Input: $A = (a_{ij})_{n \times k}$,
- A is tall matrix, $n > k$, full rank
- $A = (a_1 \; a_2 \; ... a_k)$, a_i - column of A

Reduced QR factorization, 1

- Input: $A = (a_{ij})_{n \times k}$,
- A is tall matrix, $n > k$, full rank
- $A = (a_1 \ a_2 \ ...a_k)$, a_i - column of A
- a_i is n-vector, columns - linearly independent

Reduced QR factorization, 1

- Input: $A = (a_{ij})_{n \times k}$,
- A is tall matrix, $n > k$, full rank
- $A = (a_1 \; a_2 \; ... a_k)$, a_i - column of A
- a_i is n-vector, columns - linearly independent
- Output: $Q = (q_{ij})_{n \times k}$, $R = (r_{ij})_{k \times k}$

Reduced QR factorization, 1

- Input: $A = (a_{ij})_{n \times k}$,
- A is tall matrix, $n > k$, full rank
- $A = (a_1 \; a_2 \; ...a_k)$, a_i - column of A
- a_i is n-vector, columns - linearly independent
- Output: $Q = (q_{ij})_{n \times k}$, $R = (r_{ij})_{k \times k}$
- Q is tall matrix, $n > k$

Reduced QR factorization, 1

- Input: $A = (a_{ij})_{n \times k}$,
- A is tall matrix, $n > k$, full rank
- $A = (a_1 \; a_2 \; ...a_k)$, a_i - column of A
- a_i is n-vector, columns - linearly independent
- Output: $Q = (q_{ij})_{n \times k}$, $R = (r_{ij})_{k \times k}$
- Q is tall matrix, $n > k$
- $Q = (q_1 \; q_2 \; ...q_k)$, q_i - column of Q

Reduced QR factorization, 1

- Input: $A = (a_{ij})_{n \times k}$,
- A is tall matrix, $n > k$, full rank
- $A = (a_1 \; a_2 \; ...a_k)$, a_i - column of A
- a_i is n-vector, columns - linearly independent
- Output: $Q = (q_{ij})_{n \times k}$, $R = (r_{ij})_{k \times k}$
- Q is tall matrix, $n > k$
- $Q = (q_1 \; q_2 \; ...q_k)$, q_i - column of Q
- q_i is n-vector

Reduced QR factorization, 1

- Input: $A = (a_{ij})_{n \times k}$,
- A is tall matrix, $n > k$, full rank
- $A = (a_1 \ a_2 \ ... a_k)$, a_i - column of A
- a_i is n-vector, columns - linearly independent
- Output: $Q = (q_{ij})_{n \times k}$, $R = (r_{ij})_{k \times k}$
- Q is tall matrix, $n > k$
- $Q = (q_1 \ q_2 \ ... q_k)$, q_i - column of Q
- q_i is n-vector
- $q_i \perp q_j, i \neq j$, $\|q_i\| = 1, i, j = 1, 2, ..., k \Rightarrow Q^T Q = I$,

Reduced QR factorization, 1

- Input: $A = (a_{ij})_{n \times k}$,
- A is tall matrix, $n > k$, full rank
- $A = (a_1 \ a_2 \ ... a_k)$, a_i - column of A
- a_i is n-vector, columns - linearly independent
- Output: $Q = (q_{ij})_{n \times k}$, $R = (r_{ij})_{k \times k}$
- Q is tall matrix, $n > k$
- $Q = (q_1 \ q_2 \ ... q_k)$, q_i - column of Q
- q_i is n-vector
- $q_i \perp q_j, i \neq j$, $\|q_i\| = 1, i, j = 1, 2, ..., k \Rightarrow Q^T Q = I$,
- R is upper triangular matrix

Reduced QR factorization, 1

- Input: $A = (a_{ij})_{n\times k}$,
- A is tall matrix, $n > k$, full rank
- $A = (a_1 \ a_2 \ ...a_k)$, a_i - column of A
- a_i is n-vector, columns - linearly independent
- Output: $Q = (q_{ij})_{n\times k}$, $R = (r_{ij})_{k\times k}$
- Q is tall matrix, $n > k$
- $Q = (q_1 \ q_2 \ ...q_k)$, q_i - column of Q
- q_i is n-vector
- $q_i \perp q_j, i \neq j$, $\|q_i\| = 1, i, j = 1, 2, ..., k \Rightarrow Q^T Q = I$,
- R is upper triangular matrix
- $$A = QR$$

Reduced QR factorization, 2

- Applying Gramm-Schmidt to $a_1 \ a_2 \ ...a_k$ yields $q_1 \ q_2 \ ...q_k$

Reduced QR factorization, 2

- Applying Gramm-Schmidt to $a_1 \ a_2 \ ...a_k$ yields $q_1 \ q_2 \ ...q_k$
- Gramm-Schmidt algorithm

Reduced QR factorization, 2

- Applying Gramm-Schmidt to $a_1\ a_2\ ...a_k$ yields $q_1\ q_2\ ...q_k$
- Gramm-Schmidt algorithm
  - Input: n-vectors $x_1, x_2, ..., x_m$

Reduced QR factorization, 2

- Applying Gramm-Schmidt to $a_1 \; a_2 \; ...a_k$ yields $q_1 \; q_2 \; ...q_k$
- Gramm-Schmidt algorithm
  - Input: n-vectors $x_1, x_2, ..., x_m$
  - Output:n-vectors $y_1, y_2, ..., y_m$

Reduced QR factorization, 2

- Applying Gramm-Schmidt to $a_1\ a_2\ ...a_k$ yields $q_1\ q_2\ ...q_k$
- Gramm-Schmidt algorithm
  - Input: n-vectors $x_1, x_2, ..., x_m$
  - Output:n-vectors $y_1, y_2, ..., y_m$
  - $y_i \perp y_j,\ i \neq j,\ i, j = 1, 2, ..., m$, inner product $(.,.)$

Reduced QR factorization, 2

- Applying Gramm-Schmidt to $a_1\ a_2\ ...a_k$ yields $q_1\ q_2\ ...q_k$
- Gramm-Schmidt algorithm
  - Input: n-vectors $x_1, x_2, ..., x_m$
  - Output:n-vectors $y_1, y_2, ..., y_m$
  - $y_i \perp y_j,\ i \neq j,\ i, j = 1, 2, ..., m$, inner product $(., .)$
  - for $i = 1, 2, ..., m$ do:

Reduced QR factorization, 2

- Applying Gramm-Schmidt to $a_1\ a_2\ \ldots a_k$ yields $q_1\ q_2\ \ldots q_k$
- Gramm-Schmidt algorithm
  - Input: n-vectors $x_1, x_2, ..., x_m$
  - Output: n-vectors $y_1, y_2, ..., y_m$
  - $y_i \perp y_j,\ i \neq j,\ i, j = 1, 2, ..., m$, inner product $(.,.)$
  - for $i = 1, 2, ..., m$ do:
    1. $\tilde{y}_i = x_i - \sum_{j=1}^{i-1}(y_j, x_i)y_j$

Reduced QR factorization, 2

- Applying Gramm-Schmidt to $a_1\ a_2\ ...a_k$ yields $q_1\ q_2\ ...q_k$
- Gramm-Schmidt algorithm
  - Input: n-vectors $x_1, x_2, ..., x_m$
  - Output:n-vectors $y_1, y_2, ..., y_m$
  - $y_i \perp y_j,\ i \neq j,\ i, j = 1, 2, ..., m$, inner product $(.,.)$
  - for $i = 1, 2, ..., m$ do:
    1. $\tilde{y}_i = x_i - \sum_{j=1}^{i-1} (y_j, x_i) y_j$
    2. if $\tilde{y}_i = 0$, terminate

Reduced QR factorization, 2

- Applying Gramm-Schmidt to $a_1\ a_2\ ...a_k$ yields $q_1\ q_2\ ...q_k$
- Gramm-Schmidt algorithm
  - Input: n-vectors $x_1, x_2, ..., x_m$
  - Output:n-vectors $y_1, y_2, ..., y_m$
  - $y_i \perp y_j,\ i \neq j,\ i, j = 1, 2, ..., m$, inner product $(.,.)$
  - for $i = 1, 2, ..., m$ do:
    1. $\tilde{y}_i = x_i - \sum_{j=1}^{i-1} (y_j, x_i) y_j$
    2. if $\tilde{y}_i = 0$, terminate
    3. $y_i = \frac{\tilde{y}_i}{\|\tilde{y}_i\|}$

Reduced QR factorization, 2

- Applying Gramm-Schmidt to $a_1\ a_2\ ...a_k$ yields $q_1\ q_2\ ...q_k$
- Gramm-Schmidt algorithm
  - Input: n-vectors $x_1, x_2, ..., x_m$
  - Output:n-vectors $y_1, y_2, ..., y_m$
  - $y_i \perp y_j,\ i \neq j,\ i, j = 1, 2, ..., m$, inner product $(., .)$
  - for $i = 1, 2, ..., m$ do:
    1. $\tilde{y}_i = x_i - \sum_{j=1}^{i-1}(y_j, x_i)y_j$
    2. if $\tilde{y}_i = 0$, terminate
    3. $y_i = \frac{\tilde{y}_i}{\|\tilde{y}_i\|}$
- $\tilde{q}_i = a_i - \sum_{j=1}^{i-1}(q_j, a_i)q_j$

Reduced QR factorization, 2

- Applying Gramm-Schmidt to $a_1\ a_2\ ...a_k$ yields $q_1\ q_2\ ...q_k$
- Gramm-Schmidt algorithm
  - Input: n-vectors $x_1, x_2, ..., x_m$
  - Output:n-vectors $y_1, y_2, ..., y_m$
  - $y_i \perp y_j,\ i \neq j,\ i, j = 1, 2, ..., m$, inner product $(., .)$
  - for $i = 1, 2, ..., m$ do:
    1. $\tilde{y}_i = x_i - \sum_{j=1}^{i-1}(y_j, x_i)y_j$
    2. if $\tilde{y}_i = 0$, terminate
    3. $y_i = \frac{\tilde{y}_i}{\|\tilde{y}_i\|}$
- $\tilde{q}_i = a_i - \sum_{j=1}^{i-1}(q_j, a_i)q_j$
- $q_i = \frac{\tilde{q}_i}{\|\tilde{q}_i\|}$

Reduced QR factorization, 2

- Applying Gramm-Schmidt to $a_1\ a_2\ ...a_k$ yields $q_1\ q_2\ ...q_k$
- Gramm-Schmidt algorithm
  - Input: n-vectors $x_1, x_2, ..., x_m$
  - Output:n-vectors $y_1, y_2, ..., y_m$
  - $y_i \perp y_j,\ i \neq j,\ i, j = 1, 2, ..., m$, inner product $(.,.)$
  - for $i = 1, 2, ..., m$ do:
    1. $\tilde{y}_i = x_i - \sum_{j=1}^{i-1} (y_j, x_i) y_j$
    2. if $\tilde{y}_i = 0$, terminate
    3. $y_i = \frac{\tilde{y}_i}{\|\tilde{y}_i\|}$
- $\tilde{q}_i = a_i - \sum_{j=1}^{i-1} (q_j, a_i) q_j$
- $q_i = \frac{\tilde{q}_i}{\|\tilde{q}_i\|}$
- $a_i = \sum_{j=1}^{i-1} (q_j, a_i) q_j + \|\tilde{q}_i\| q_i$

Reduced QR factorization, 2

- Applying Gramm-Schmidt to $a_1\ a_2\ ...a_k$ yields $q_1\ q_2\ ...q_k$
- Gramm-Schmidt algorithm
  - Input: n-vectors $x_1, x_2, ..., x_m$
  - Output:n-vectors $y_1, y_2, ..., y_m$
  - $y_i \perp y_j,\ i \neq j,\ i, j = 1, 2, ..., m$, inner product $(.,.)$
  - for $i = 1, 2, ..., m$ do:
    1. $\tilde{y}_i = x_i - \sum_{j=1}^{i-1}(y_j, x_i)y_j$
    2. if $\tilde{y}_i = 0$, terminate
    3. $y_i = \frac{\tilde{y}_i}{\|\tilde{y}_i\|}$
- $\tilde{q}_i = a_i - \sum_{j=1}^{i-1}(q_j, a_i)q_j$
- $q_i = \frac{\tilde{q}_i}{\|\tilde{q}_i\|}$
- $a_i = \sum_{j=1}^{i-1}(q_j, a_i)q_j + \|\tilde{q}_i\| q_i$
- $r_{ij} = \begin{cases} i < j : r_{ij} = (a_j, q_i) \\ i = j : r_{ii} = \|\tilde{q}_i\| \\ i > j : r_{ij} = 0 \end{cases}$

Reduced QR factorization, 3

- $a_i = \sum_{j=1}^{i-1} (q_j, a_i) q_j + \|\tilde{q}_i\| q_i$
  - $r_{ij} = \begin{cases} i < j : r_{ij} = (a_j, q_i) \\ i = j : r_{ii} = \|\tilde{q}_i\| \\ i > j : r_{ij} = 0 \end{cases}$,

Reduced QR factorization, 3

- $a_i = \sum_{j=1}^{i-1}(q_j, a_i)q_j + \|\tilde{q}_i\| q_i$
  - $$r_{ij} = \begin{cases} i < j : r_{ij} = (a_j, q_i) \\ i = j : r_{ii} = \|\tilde{q}_i\| \\ i > j : r_{ij} = 0 \end{cases}, \quad A = QR$$

Reduced QR factorization, 3

- $a_i = \sum_{j=1}^{i-1}(q_j, a_i)q_j + \|\tilde{q}_i\| q_i$
- $r_{ij} = \begin{cases} i < j : r_{ij} = (a_j, q_i) \\ i = j : r_{ii} = \|\tilde{q}_i\| \\ i > j : r_{ij} = 0 \end{cases}, A = QR$
-

$$a_1 = r_{11} q_1 = \begin{pmatrix} q_1 & q_2 & \dots & q_k \end{pmatrix} \begin{pmatrix} r_{11} \\ 0 \\ \dots \\ 0 \end{pmatrix}$$

Reduced QR factorization, 3

- $a_i = \sum_{j=1}^{i-1}(q_j, a_i)q_j + \|\tilde{q}_i\| q_i$
- $r_{ij} = \begin{cases} i < j : r_{ij} = (a_j, q_i) \\ i = j : r_{ii} = \|\tilde{q}_i\| \\ i > j : r_{ij} = 0 \end{cases}, \quad A = QR$
-
$$a_1 = r_{11}q_1 = \begin{pmatrix} q_1 & q_2 & \dots & q_k \end{pmatrix} \begin{pmatrix} r_{11} \\ 0 \\ \dots \\ 0 \end{pmatrix}$$
-
$$a_2 = r_{12}q_1 + r_{22}q_2 = \begin{pmatrix} q_1 & q_2 & \dots & q_k \end{pmatrix} \begin{pmatrix} r_{12} \\ r_{22} \\ \dots \\ 0 \end{pmatrix}$$

Reduced QR factorization, 3

- $a_i = \sum_{j=1}^{i-1}(q_j, a_i)q_j + \|\tilde{q}_i\|q_i$
- $r_{ij} = \begin{cases} i < j : r_{ij} = (a_j, q_i) \\ i = j : r_{ii} = \|\tilde{q}_i\| \\ i > j : r_{ij} = 0 \end{cases}, \quad A = QR$
- $$a_1 = r_{11}q_1 = \begin{pmatrix} q_1 & q_2 & \dots & q_k \end{pmatrix} \begin{pmatrix} r_{11} \\ 0 \\ \dots \\ 0 \end{pmatrix}$$
- $$a_2 = r_{12}q_1 + r_{22}q_2 = \begin{pmatrix} q_1 & q_2 & \dots & q_k \end{pmatrix} \begin{pmatrix} r_{12} \\ r_{22} \\ \dots \\ 0 \end{pmatrix}$$
- $$a_k = r_{1k}q_1 + r_{2k}q_2 + \dots + r_{kk}q_k = \begin{pmatrix} q_1 & q_2 & \dots & q_k \end{pmatrix} \begin{pmatrix} r_{1k} \\ r_{2k} \\ \dots \\ r_{kk} \end{pmatrix}$$

Reduced QR factorization, 4

- $$a_1 = r_{11} q_1 = \begin{pmatrix} q_1 & q_2 & \dots & q_k \end{pmatrix} \begin{pmatrix} r_{11} \\ 0 \\ \dots \\ 0 \end{pmatrix}$$
- $$a_2 = r_{12} q_1 + r_{22} q_2 = \begin{pmatrix} q_1 & q_2 & \dots & q_k \end{pmatrix} \begin{pmatrix} r_{12} \\ r_{22} \\ \dots \\ 0 \end{pmatrix}$$
- $$a_k = r_{1k} q_1 + r_{2k} q_2 + \dots + r_{kk} q_k = \begin{pmatrix} q_1 & q_2 & \dots & q_k \end{pmatrix} \begin{pmatrix} r_{1k} \\ r_{2k} \\ \dots \\ r_{kk} \end{pmatrix}$$

Reduced QR factorization, 4

- $$a_1 = r_{11} q_1 = \begin{pmatrix} q_1 & q_2 & \dots & q_k \end{pmatrix} \begin{pmatrix} r_{11} \\ 0 \\ \dots \\ 0 \end{pmatrix}$$
- $$a_2 = r_{12} q_1 + r_{22} q_2 = \begin{pmatrix} q_1 & q_2 & \dots & q_k \end{pmatrix} \begin{pmatrix} r_{12} \\ r_{22} \\ \dots \\ 0 \end{pmatrix}$$
- $$a_k = r_{1k} q_1 + r_{2k} q_2 + \dots + r_{kk} q_k = \begin{pmatrix} q_1 & q_2 & \dots & q_k \end{pmatrix} \begin{pmatrix} r_{1k} \\ r_{2k} \\ \dots \\ r_{kk} \end{pmatrix}$$
- $$\begin{pmatrix} a_1 & a_2 & \dots & a_k \end{pmatrix} = \begin{pmatrix} q_1 & q_2 & \dots & q_k \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1k} \\ 0 & r_{22} & \dots & r_{1k} \\ . & . & \dots & . \\ 0 & 0 & \dots & r_{kk} \end{pmatrix}$$

Reduced QR factorization, 5

- $$\begin{pmatrix} a_1 & a_2 & \dots & a_k \end{pmatrix} = \begin{pmatrix} q_1 & q_2 & \dots & q_k \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1k} \\ 0 & r_{22} & \dots & r_{1k} \\ . & . & \dots & . \\ 0 & 0 & \dots & r_{kk} \end{pmatrix}$$
- $$A = QR$$
- Q: Is reduced QR factorization always possible?

Reduced QR factorization, 5

- $$\begin{pmatrix} a_1 & a_2 & \dots & a_k \end{pmatrix} = \begin{pmatrix} q_1 & q_2 & \dots & q_k \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1k} \\ 0 & r_{22} & \dots & r_{1k} \\ . & . & \dots & . \\ 0 & 0 & \dots & r_{kk} \end{pmatrix}$$
- $$A = QR$$
- Q: Is reduced QR factorization always possible?
- A:

Theorem 13.5

For $A \in \mathcal{R}^{n\times m}$ of full rank always exist $Q \in \mathcal{R}^{n\times m}$ with orthonormal columns and upper triangular $R \in \mathcal{R}^{m\times m}$ such that $A = QR$

Reduced QR factorization

```
A= [[ 1  2  3]
 [ 4  5  6]
 [ 7  8  9]
 [10 11 12]]
Q= [[-0.07761505 -0.83305216  0.53358462]
 [-0.31046021 -0.45123659 -0.8036038 ]
 [-0.54330537 -0.06942101  0.00645373]
 [-0.77615053  0.31239456  0.26356544]]
R= [[-1.28840987e+01 -1.45916299e+01 -1.62991610e+01]
 [ 0.00000000e+00 -1.04131520e+00 -2.08263040e+00]
 [ 0.00000000e+00  0.00000000e+00 -3.39618744e-15]]
A-QR= [[-1.33226763e-15 -7.10542736e-15 -5.32907052e-15]
 [-8.88178420e-16 -1.77635684e-15 -8.88178420e-16]
 [ 0.00000000e+00 -3.55271368e-15 -3.55271368e-15]
 [-1.77635684e-15 -5.32907052e-15 -3.55271368e-15]]
Pseudo inverse =  [[-0.48333333 -0.24444444 -0.00555556  0.23333333]
 [-0.03333333 -0.01111111  0.01111111  0.03333333]
 [ 0.41666667  0.22222222  0.02777778 -0.16666667]]
```

Figure: Pseudo inverse, $A^{\dagger} = R^{-1}Q^{T}$

Linear Least Squares

Problem 13.6

Linear Least Squares Problem

Linear Least Squares

Problem 13.6

Linear Least Squares Problem

- $A \in \mathbb{R}^{n \times m}, n > m, x \in \mathbb{R}^m, b \in \mathbb{R}^n$

Linear Least Squares

Problem 13.6

Linear Least Squares Problem

- $A \in \mathbb{R}^{n \times m}, n > m, x \in \mathbb{R}^m, b \in \mathbb{R}^n$
- $Ax = b$ - over determined system

Linear Least Squares

Problem 13.6

Linear Least Squares Problem

- $A \in \mathbb{R}^{n \times m}, n > m, x \in \mathbb{R}^m, b \in \mathbb{R}^n$
- $Ax = b$ - over determined system
- $f(x) = \|Ax - b\|_2$

Linear Least Squares

Problem 13.6

Linear Least Squares Problem

- $A \in \mathbb{R}^{n \times m}, n > m, x \in \mathbb{R}^m, b \in \mathbb{R}^n$
- $Ax = b$ - over determined system
- $f(x) = \|Ax - b\|_2$
- $x_* = arg \min f(x)$

Linear Least Squares

Problem 13.6

Linear Least Squares Problem

- $A \in \mathbb{R}^{n \times m}, n > m, x \in \mathbb{R}^m, b \in \mathbb{R}^n$
- $Ax = b$ - over determined system
- $f(x) = \|Ax - b\|_2$
- $x_* = arg \min f(x)$

Definition 13.7

- If the least squares problem has more than one solution

Linear Least Squares

Problem 13.6

Linear Least Squares Problem

- $A \in \mathbb{R}^{n \times m}, n > m, x \in \mathbb{R}^m, b \in \mathbb{R}^n$
- $Ax = b$ - over determined system
- $f(x) = \|Ax - b\|_2$
- $x_* = arg \min f(x)$

Definition 13.7

- If the least squares problem has more than one solution
- minimal-length solution minimal-norm solution $=$ solution with least Eucledean norm

Linear Least Squares

Solution of Linear Least Squares Problem

- $x_* = arg \min f(x) \Rightarrow \nabla f(x) = 0$

Linear Least Squares

Solution of Linear Least Squares Problem

- $x_* = arg\min f(x) \Rightarrow \nabla f(x) = 0$

$$\Downarrow$$

Linear Least Squares

Solution of Linear Least Squares Problem

- $x_* = arg \min f(x) \Rightarrow \nabla f(x) = 0$

$$\Downarrow$$

- $A^T A x = A^T b$

Linear Least Squares

Solution of Linear Least Squares Problem

- $x_* = arg\min f(x) \Rightarrow \nabla f(x) = 0$

$$\Downarrow$$

- $A^T Ax = A^T b$ - Normal equation

Linear Least Squares

Solution of Linear Least Squares Problem

- $x_* = arg\min f(x) \Rightarrow \nabla f(x) = 0$

$$\Downarrow$$

- $A^T A x = A^T b$ - Normal equation
- $\exists (A^T A)^{-1}$,

Linear Least Squares

Solution of Linear Least Squares Problem

- $x_* = arg\min f(x) \Rightarrow \nabla f(x) = 0$

$$\Downarrow$$

- $A^T A x = A^T b$ - Normal equation
- $\exists (A^T A)^{-1}, \Rightarrow x = (A^T A)^{-1} A^T b$

Linear Least Squares

Solution of Linear Least Squares Problem

- $x_* = arg \min f(x) \Rightarrow \nabla f(x) = 0$

$$\Downarrow$$

- $A^T A x = A^T b$ - Normal equation
- $\exists (A^T A)^{-1}, \Rightarrow x = (A^T A)^{-1} A^T b$

Definition 13.8

Pseudo inverse of a matrix

- $A \in \mathbb{R}^{n \times m}, n > m$

Linear Least Squares

Solution of Linear Least Squares Problem

- $x_* = arg\min f(x) \Rightarrow \nabla f(x) = 0$

$$\Downarrow$$

- $A^T A x = A^T b$ - Normal equation
- $\exists (A^T A)^{-1}, \Rightarrow x = (A^T A)^{-1} A^T b$

Definition 13.8

Pseudo inverse of a matrix

- $A \in \mathbb{R}^{n \times m}, n > m$
- $\exists (A^T A)^{-1}$

Linear Least Squares

Solution of Linear Least Squares Problem

- $x_* = arg\min f(x) \Rightarrow \nabla f(x) = 0$

$$\Downarrow$$

- $A^T A x = A^T b$ - Normal equation
- $\exists (A^T A)^{-1}, \Rightarrow x = (A^T A)^{-1} A^T b$

Definition 13.8

Pseudo inverse of a matrix

- $A \in \mathbb{R}^{n \times m}, n > m$
- $\exists (A^T A)^{-1}$
- $A^\dagger = (A^T A)^{-1} A^T$

Linear Least Squares

Pseudo inverse and QR factorization

Linear Least Squares

Pseudo inverse and QR factorization

- $A \in \mathbb{R}^{n\times m}, n > m$

Linear Least Squares

Pseudo inverse and QR factorization

- $A \in \mathbb{R}^{n\times m}, n > m$
- $rank(A) = m$, $A = QR$

Linear Least Squares

Pseudo inverse and QR factorization

- $A \in \mathbb{R}^{n \times m}, n > m$
- $rank(A) = m$, $A = QR$
- $A^{\dagger} = (A^T A)^{-1} A^T$

Linear Least Squares

Pseudo inverse and QR factorization

- $A \in \mathbb{R}^{n \times m}, n > m$
- $rank(A) = m$, $A = QR$
- $A^{\dagger} = (A^T A)^{-1} A^T$
- $A^T = (QR)^T = R^T Q^T$,

Linear Least Squares

Pseudo inverse and QR factorization

- $A \in \mathbb{R}^{n \times m}, n > m$
- $rank(A) = m$, $A = QR$
- $A^{\dagger} = (A^T A)^{-1} A^T$
- $A^T = (QR)^T = R^T Q^T$, $A^T A = R^T Q^T Q R = R^T R$

Linear Least Squares

Pseudo inverse and QR factorization

- $A \in \mathbb{R}^{n \times m}, n > m$
- $rank(A) = m$, $A = QR$
- $A^\dagger = (A^T A)^{-1} A^T$
- $A^T = (QR)^T = R^T Q^T$, $A^T A = R^T Q^T QR = R^T R$
- $(A^T A)^{-1} = (R^T R)^{-1} = R^{-1} R^{-T}$

Linear Least Squares

Pseudo inverse and QR factorization

- $A \in \mathbb{R}^{n \times m}, n > m$
- $rank(A) = m$, $A = QR$
- $A^{\dagger} = (A^T A)^{-1} A^T$
- $A^T = (QR)^T = R^T Q^T$, $A^T A = R^T Q^T Q R = R^T R$
- $(A^T A)^{-1} = (R^T R)^{-1} = R^{-1} R^{-T}$
- $A^{\dagger} = (A^T A)^{-1} A^T = R^{-1} R^{-T} R^T Q^T$

Linear Least Squares

Pseudo inverse and QR factorization

- $A \in \mathbb{R}^{n \times m}, n > m$
- $rank(A) = m$, $A = QR$
- $A^{\dagger} = (A^T A)^{-1} A^T$
- $A^T = (QR)^T = R^T Q^T$, $A^T A = R^T Q^T Q R = R^T R$
- $(A^T A)^{-1} = (R^T R)^{-1} = R^{-1} R^{-T}$
- $A^{\dagger} = (A^T A)^{-1} A^T = R^{-1} R^{-T} R^T Q^T$
- $A^{\dagger} = R^{-1} Q^T$

Linear Least Squares

Pseudo inverse and QR factorization

- $A \in \mathbb{R}^{n \times m}, n > m$
- $rank(A) = m$, $A = QR$
- $A^{\dagger} = (A^T A)^{-1} A^T$
- $A^T = (QR)^T = R^T Q^T$, $A^T A = R^T Q^T QR = R^T R$
- $(A^T A)^{-1} = (R^T R)^{-1} = R^{-1} R^{-T}$
- $A^{\dagger} = (A^T A)^{-1} A^T = R^{-1} R^{-T} R^T Q^T$
- $A^{\dagger} = R^{-1} Q^T$

Least squares solution and QR factorization

Linear Least Squares

Pseudo inverse and QR factorization

- $A \in \mathbb{R}^{n \times m}, n > m$
- $rank(A) = m$, $A = QR$
- $A^{\dagger} = (A^T A)^{-1} A^T$
- $A^T = (QR)^T = R^T Q^T$, $A^T A = R^T Q^T Q R = R^T R$
- $(A^T A)^{-1} = (R^T R)^{-1} = R^{-1} R^{-T}$
- $A^{\dagger} = (A^T A)^{-1} A^T = R^{-1} R^{-T} R^T Q^T$
- $A^{\dagger} = R^{-1} Q^T$

Least squares solution and QR factorization

- $A \in \mathbb{R}^{n \times m}, n > m, x \in \mathbb{R}^m, b \in \mathbb{R}^n$

Linear Least Squares

Pseudo inverse and QR factorization

- $A \in \mathbb{R}^{n \times m}, n > m$
- $rank(A) = m$, $A = QR$
- $A^{\dagger} = (A^T A)^{-1} A^T$
- $A^T = (QR)^T = R^T Q^T$, $A^T A = R^T Q^T Q R = R^T R$
- $(A^T A)^{-1} = (R^T R)^{-1} = R^{-1} R^{-T}$
- $A^{\dagger} = (A^T A)^{-1} A^T = R^{-1} R^{-T} R^T Q^T$
- $A^{\dagger} = R^{-1} Q^T$

Least squares solution and QR factorization

- $A \in \mathbb{R}^{n \times m}, n > m, x \in \mathbb{R}^m, b \in \mathbb{R}^n$
- $rank(A) = m$, $A = QR$

Linear Least Squares

Pseudo inverse and QR factorization

- $A \in \mathbb{R}^{n \times m}, n > m$
- $rank(A) = m$, $A = QR$
- $A^{\dagger} = (A^T A)^{-1} A^T$
- $A^T = (QR)^T = R^T Q^T$, $A^T A = R^T Q^T Q R = R^T R$
- $(A^T A)^{-1} = (R^T R)^{-1} = R^{-1} R^{-T}$
- $A^{\dagger} = (A^T A)^{-1} A^T = R^{-1} R^{-T} R^T Q^T$
- $A^{\dagger} = R^{-1} Q^T$

Least squares solution and QR factorization

- $A \in \mathbb{R}^{n \times m}, n > m, x \in \mathbb{R}^m, b \in \mathbb{R}^n$
- $rank(A) = m$, $A = QR$
- $Ax = b$

Linear Least Squares

Pseudo inverse and QR factorization

- $A \in \mathbb{R}^{n\times m}, n > m$
- $rank(A) = m$, $A = QR$
- $A^\dagger = (A^TA)^{-1}A^T$
- $A^T = (QR)^T = R^TQ^T$, $A^TA = R^TQ^TQR = R^TR$
- $(A^TA)^{-1} = (R^TR)^{-1} = R^{-1}R^{-T}$
- $A^\dagger = (A^TA)^{-1}A^T = R^{-1}R^{-T}R^TQ^T$
- $A^\dagger = R^{-1}Q^T$

Least squares solution and QR factorization

- $A \in \mathbb{R}^{n\times m}, n > m, x \in \mathbb{R}^m, b \in \mathbb{R}^n$
- $rank(A) = m$, $A = QR$
- $Ax = b$
- $c = Q^Tb, Rx = c, x = R^{-1}c$

Linear Least Squares

Least squares solution and QR factorization

- $A \in \mathbb{R}^{n \times m}, n > m$
- $rank(A) = m$, $A = QR$
- $Ax = b$
- $c = Q^T b, Rx = c, x = R^{-1} c$

- Rank deficient case: $rank(A) < m$

Linear Least Squares

Least squares solution and QR factorization

- $A \in \mathbb{R}^{n \times m}, n > m$
- $rank(A) = m$, $A = QR$
- $Ax = b$
- $c = Q^T b, Rx = c, x = R^{-1}c$

- Rank deficient case: $rank(A) < m$
- Solution is not unique

Linear Least Squares

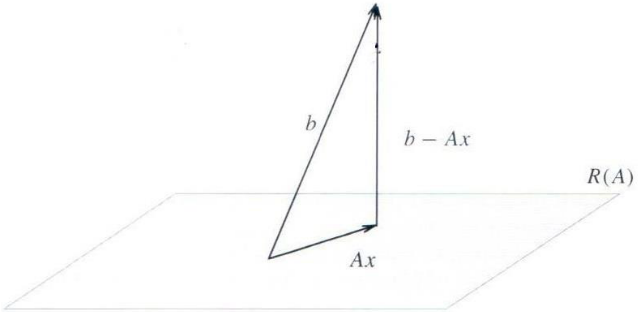


Figure: Geometric interpretation of Least Squares, Biswa Natt Datta

- $A \in \mathcal{R}^{n \times m}, R(A) = \{y : y = Ax, \forall x \in \mathcal{R}^m\}$

Linear Least Squares

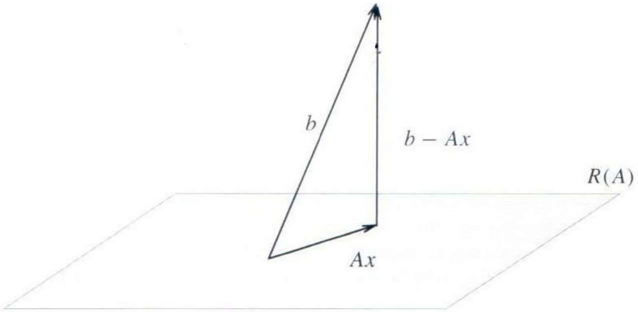


Figure: Geometric interpretation of Least Squares, Biswa Natt Datta

- $A \in \mathcal{R}^{n\times m}, R(A) = \{y : y = Ax, \forall x \in \mathcal{R}^m\}$
- $x_* = arg\min_{x\in\mathcal{R}^m} \|Ax - b\|_2$

Linear Least Squares

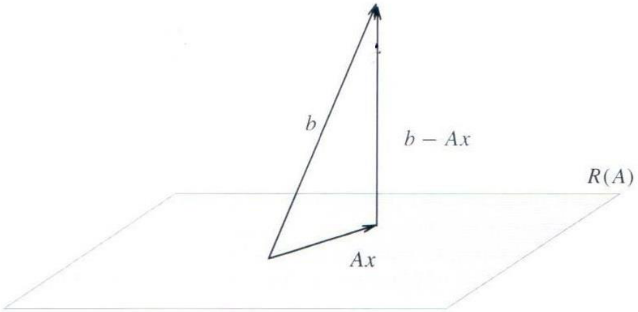


Figure: Geometric interpretation of Least Squares, Biswa Natt Datta

- $A \in \mathcal{R}^{n\times m}, R(A) = \{y : y = Ax, \forall x \in \mathcal{R}^m\}$
- $x_* = arg\min_{x\in\mathcal{R}^m} \|Ax - b\|_2$
- $x_* \in R(A)$, minimal distance to b

Linear Least Squares

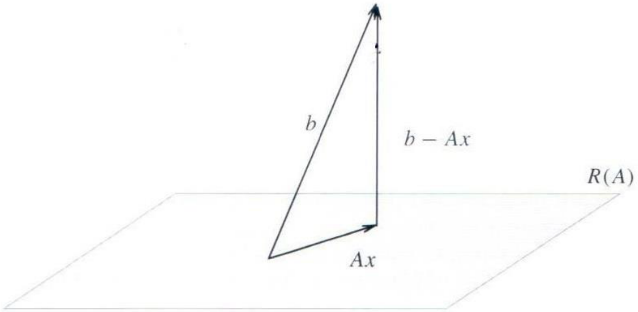


Figure: Geometric interpretation of Least Squares, Biswa Natt Datta

- $A \in \mathcal{R}^{n\times m}, R(A) = \{y : y = Ax, \forall x \in \mathcal{R}^m\}$
- $x_* = arg\,\min_{x\in\mathcal{R}^m} \|Ax - b\|_2$
- $x_* \in R(A)$, minimal distance to b $\Downarrow$
- $(b - Ax) \perp R(A)$

Linear Least Squares

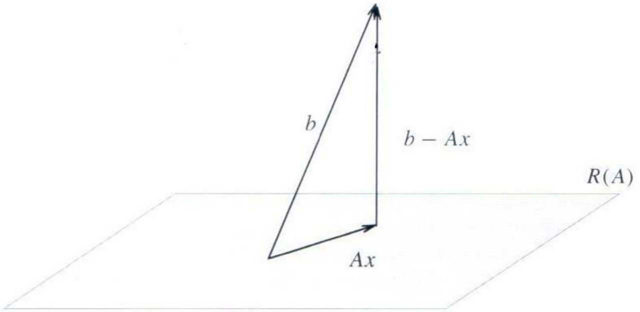


Figure: Geometric interpretation of Least Squares, Biswa Natt Datta

- $A \in \mathcal{R}^{n\times m}, R(A) = \{y : y = Ax, \forall x \in \mathcal{R}^m\}$
- $x_* = arg\min_{x\in\mathcal{R}^m} \|Ax - b\|_2$
- $x_* \in R(A)$, minimal distance to b $\Downarrow$
- $(b - Ax) \perp R(A)$
- Solution to linear least squares problem always exists

Constrained Least Squares

Problem 13.9

Constrained Least Squares Problem

Constrained Least Squares

Problem 13.9

Constrained Least Squares Problem

- $A \in \mathbb{R}^{n \times m}, n > m$

Constrained Least Squares

Problem 13.9

Constrained Least Squares Problem

- $A \in \mathbb{R}^{n \times m}, n > m$
- $C \in \mathbb{R}^{p \times m}$

Constrained Least Squares

Problem 13.9

Constrained Least Squares Problem

- $A \in \mathbb{R}^{n \times m}, n > m$
- $C \in \mathbb{R}^{p \times m}$
- $x \in \mathbb{R}^m, b \in \mathbb{R}^n, d \in \mathbb{R}^p$

Constrained Least Squares

Problem 13.9

Constrained Least Squares Problem

- $A \in \mathbb{R}^{n \times m}, n > m$
- $C \in \mathbb{R}^{p \times m}$
- $x \in \mathbb{R}^m, b \in \mathbb{R}^n, d \in \mathbb{R}^p$
- $Ax = b$ - over determined system

Constrained Least Squares

Problem 13.9

Constrained Least Squares Problem

- $A \in \mathbb{R}^{n \times m}, n > m$
- $C \in \mathbb{R}^{p \times m}$
- $x \in \mathbb{R}^m, b \in \mathbb{R}^n, d \in \mathbb{R}^p$
- $Ax = b$ - over determined system
- $f(x) = \|Ax - b\|_2$

Constrained Least Squares

Problem 13.9

Constrained Least Squares Problem

- $A \in \mathbb{R}^{n \times m}, n > m$
- $C \in \mathbb{R}^{p \times m}$
- $x \in \mathbb{R}^m, b \in \mathbb{R}^n, d \in \mathbb{R}^p$
- $Ax = b$ - over determined system
- $f(x) = \|Ax - b\|_2$
- $x_* = arg\min_{Cx=d} f(x)$

Constrained Least Squares

Problem 13.9

Constrained Least Squares Problem

- $A \in \mathbb{R}^{n \times m}, n > m$
- $C \in \mathbb{R}^{p \times m}$
- $x \in \mathbb{R}^m, b \in \mathbb{R}^n, d \in \mathbb{R}^p$
- $Ax = b$ - over determined system
- $f(x) = \|Ax - b\|_2$
- $x_* = arg\min_{Cx=d} f(x)$

Constrains

- $Cx = d, C \in \mathbb{R}^{p \times m}, d \in \mathbb{R}^p$

Constrained Least Squares

Problem 13.9

Constrained Least Squares Problem

- $A \in \mathbb{R}^{n\times m}, n > m$
- $C \in \mathbb{R}^{p\times m}$
- $x \in \mathbb{R}^m, b \in \mathbb{R}^n, d \in \mathbb{R}^p$
- $Ax = b$ - over determined system
- $f(x) = \|Ax - b\|_2$
- $x_* = arg\min_{Cx=d} f(x)$

Constrains

- $Cx = d, C \in \mathbb{R}^{p\times m}, d \in \mathbb{R}^p$
- $C = \begin{pmatrix} c_1^T \\ c_2^T \\ \dots \\ c_p^T \end{pmatrix}, c_i^T \in \mathbb{R}^m$

Constrained Least Squares

Problem 13.9

Constrained Least Squares Problem

- $A \in \mathbb{R}^{n \times m}, n > m$
- $C \in \mathbb{R}^{p \times m}$
- $x \in \mathbb{R}^m, b \in \mathbb{R}^n, d \in \mathbb{R}^p$
- $Ax = b$ - over determined system
- $f(x) = \|Ax - b\|_2$
- $x_* = arg\min_{Cx=d} f(x)$

Constrains

- $Cx = d, C \in \mathbb{R}^{p \times m}, d \in \mathbb{R}^p$
- $C = \begin{pmatrix} c_1^T \\ c_2^T \\ \dots \\ c_p^T \end{pmatrix}, c_i^T \in \mathbb{R}^m \ c_i^T x = d_i, i = 1, 2, \dots, p, \ d = (d_1, \dots, d_p)^T$

Constrained Least Squares

Definition 13.10

$x \in \mathcal{R}^m$ is feasible if $Cx = d$

Constrained Least Squares

Definition 13.10

$x \in \mathcal{R}^m$ is feasible if $Cx = d$

Definition 13.11

$\hat{x} \in \mathcal{R}^m$ is optimal if it is feasible and if $\|A\hat{x} - b\|_2 \leq \|Ax - b\|_2 \ \ \forall x \in \mathcal{R}^m$

Constrained Least Squares

Definition 13.10

$x \in \mathcal{R}^m$ is feasible if $Cx = d$

Definition 13.11

$\hat{x} \in \mathcal{R}^m$ is optimal if it is feasible and if $\|A\hat{x} - b\|_2 \leq \|Ax - b\|_2 \quad \forall x \in \mathcal{R}^m$

Is optimal solution unique?

Constrained Least Squares

Definition 13.10

$x \in \mathcal{R}^m$ is feasible if $Cx = d$

Definition 13.11

$\hat{x} \in \mathcal{R}^m$ is optimal if it is feasible and if $\|A\hat{x} - b\|_2 \leq \|Ax - b\|_2 \;\; \forall x \in \mathcal{R}^m$

Is optimal solution unique?

Problem 13.12

Least Norm Problem

Constrained Least Squares

Definition 13.10

$x \in \mathcal{R}^m$ is feasible if $Cx = d$

Definition 13.11

$\hat{x} \in \mathcal{R}^m$ is optimal if it is feasible and if $\|A\hat{x} - b\|_2 \leq \|Ax - b\|_2 \quad \forall x \in \mathcal{R}^m$

Is optimal solution unique?

Problem 13.12

Least Norm Problem

- $C \in \mathbb{R}^{p \times m}$

Constrained Least Squares

Definition 13.10

$x \in \mathcal{R}^m$ is feasible if $Cx = d$

Definition 13.11

$\hat{x} \in \mathcal{R}^m$ is optimal if it is feasible and if $\|A\hat{x} - b\|_2 \leq \|Ax - b\|_2 \quad \forall x \in \mathcal{R}^m$

Is optimal solution unique?

Problem 13.12

Least Norm Problem

- $C \in \mathbb{R}^{p \times m}$
- $x \in \mathbb{R}^m, d \in \mathbb{R}^p$

Constrained Least Squares

Definition 13.10

$x \in \mathcal{R}^m$ is feasible if $Cx = d$

Definition 13.11

$\hat{x} \in \mathcal{R}^m$ is optimal if it is feasible and if $\|A\hat{x} - b\|_2 \leq \|Ax - b\|_2 \ \ \forall x \in \mathcal{R}^m$

Is optimal solution unique?

Problem 13.12

Least Norm Problem

- $C \in \mathbb{R}^{p \times m}$
- $x \in \mathbb{R}^m, d \in \mathbb{R}^p$
- $f(x) = \|x\|_2$

Constrained Least Squares

Definition 13.10

$x \in \mathcal{R}^m$ is feasible if $Cx = d$

Definition 13.11

$\hat{x} \in \mathcal{R}^m$ is optimal if it is feasible and if $\|A\hat{x} - b\|_2 \leq \|Ax - b\|_2 \quad \forall x \in \mathcal{R}^m$

Is optimal solution unique?

Problem 13.12

Least Norm Problem

- $C \in \mathbb{R}^{p \times m}$
- $x \in \mathbb{R}^m, d \in \mathbb{R}^p$
- $f(x) = \|x\|_2$
- $x_* = arg\,\min_{Cx=d} f(x)$

Constrained Least Squares

Lagrange multipliers

- $$L(x, \lambda) = \|Ax - b\|_2 + \sum_{i=1}^{p} \lambda_i (c_i^T x - d_i)$$

Constrained Least Squares

Lagrange multipliers

- $$L(x, \lambda) = \|Ax - b\|_2 + \sum_{i=1}^{p} \lambda_i (c_i^T x - d_i)$$
- $$L(x, \lambda) = (Ax - b, Ax - b) + (Cx - d, \lambda)$$

Constrained Least Squares

Lagrange multipliers

- $$L(x,\lambda) = \|Ax-b\|_2 + \sum_{i=1}^{p} \lambda_i (c_i^T x - d_i)$$
- $$L(x,\lambda) = (Ax-b, Ax-b) + (Cx-d, \lambda)$$

Optimality conditions

- $$\frac{\partial L}{\partial x_i} = 0, i = 1, ..., m$$

Constrained Least Squares

Lagrange multipliers

- $$L(x, \lambda) = \|Ax - b\|_2 + \sum_{i=1}^{p} \lambda_i (c_i^T x - d_i)$$
- $$L(x, \lambda) = (Ax - b, Ax - b) + (Cx - d, \lambda)$$

Optimality conditions

- $$\frac{\partial L}{\partial x_i} = 0, i = 1, ..., m$$
- $$2(A^T A x - A^T b) + C^T \lambda = 0$$

Constrained Least Squares

Lagrange multipliers

- $$L(x,\lambda) = \|Ax-b\|_2 + \sum_{i=1}^{p} \lambda_i(c_i^T x - d_i)$$
- $$L(x,\lambda) = (Ax-b, Ax-b) + (Cx-d, \lambda)$$

Optimality conditions

- $$\frac{\partial L}{\partial x_i} = 0, i = 1, ..., m$$
- $$2(A^T Ax - A^T b) + C^T \lambda = 0$$
- $$\frac{\partial L}{\partial \lambda_j} = 0, j = 1, ..., p$$

Constrained Least Squares

Lagrange multipliers

- $$L(x,\lambda) = \|Ax - b\|_2 + \sum_{i=1}^{p} \lambda_i (c_i^T x - d_i)$$
- $$L(x,\lambda) = (Ax - b, Ax - b) + (Cx - d, \lambda)$$

Optimality conditions

- $$\frac{\partial L}{\partial x_i} = 0, i = 1, ..., m$$
- $$2(A^T A x - A^T b) + C^T \lambda = 0$$
- $$\frac{\partial L}{\partial \lambda_j} = 0, j = 1, ..., p$$
- $$Cx = d$$

Constrained Least Squares

Lagrangian function, necessary conditions of minimum

$$\begin{aligned} 2(A^T Ax - A^T b) + C^T \lambda &= 0 \\ Cx &= d \end{aligned}$$

Constrained Least Squares

Lagrangian function, necessary conditions of minimum

$$\begin{aligned} 2(A^T A x - A^T b) + C^T \lambda &= 0 \\ Cx &= d \end{aligned}$$

$$\begin{pmatrix} 2A^T A & C^T \\ C & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} 2A^T b \\ d \end{pmatrix}$$

Constrained Least Squares

Lagrangian function, necessary conditions of minimum

$$
\begin{aligned}
2(A^T Ax - A^T b) + C^T \lambda = 0 \\
Cx = d
\end{aligned}
$$

$$
\begin{pmatrix} 2A^T A & C^T \\ C & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} 2A^T b \\ d \end{pmatrix}
$$

KKT equation (William Karush, Harold Kuhn and Albert Tucker)

Constrained Least Squares

Lagrangian function, necessary conditions of minimum

$$
\begin{aligned}
2(A^T Ax - A^T b) + C^T\lambda &= 0 \\
Cx &= d
\end{aligned}
$$

$$
\begin{pmatrix} 2A^T A & C^T \\ C & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} 2A^T b \\ d \end{pmatrix}
$$

KKT equation (William Karush, Harold Kuhn and Albert Tucker)

Theorem 13.13

KKT matrix is invertible iff

Constrained Least Squares

Lagrangian function, necessary conditions of minimum

$$
\begin{aligned}
2(A^T A x - A^T b) + C^T \lambda &= 0 \\
Cx &= d
\end{aligned}
$$

$$
\begin{pmatrix} 2A^T A & C^T \\ C & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} 2A^T b \\ d \end{pmatrix}
$$

KKT equation (William Karush, Harold Kuhn and Albert Tucker)

Theorem 13.13

KKT matrix is invertible iff

- *C has linearly independent rows*

Constrained Least Squares

Lagrangian function, necessary conditions of minimum

$$
\begin{aligned}
2(A^T A x - A^T b) + C^T \lambda &= 0 \\
Cx &= d
\end{aligned}
$$

$$
\begin{pmatrix} 2A^T A & C^T \\ C & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} 2A^T b \\ d \end{pmatrix}
$$

KKT equation (William Karush, Harold Kuhn and Albert Tucker)

Theorem 13.13

KKT matrix is invertible iff

- *C has linearly independent rows*
- *$\begin{pmatrix} A \\ C \end{pmatrix}$ has linearly independent columns*

Q & A