Least Squares Problems

Definition

Consider the possibly overdetermined system of equations

$$Ax = b \tag{1}$$

where $A \in \mathbb{R}^{n \times m}$ and a vector $b \in \mathbb{R}^n$, $n \geq m$. It may not have an exact solution if $n > m$ or A is square but singular. The **Least Squares Problem** associated with this system is about finding $x_0 \in \mathbb{R}^m$ such that

$$\|b - Ax_0\|_2 = \min_{x \in \mathbb{R}^m} \|b - Ax\|_2.$$

Definition

Consider the possibly overdetermined system of equations

$$Ax = b \tag{1}$$

where $A \in \mathbb{R}^{n\times m}$ and a vector $b \in \mathbb{R}^n$, $n \geq m$. It may not have an exact solution if $n > m$ or A is square but singular. The **Least Squares Problem** associated with this system is about finding $x_0 \in \mathbb{R}^m$ such that

$$\|b - Ax_0\|_2 = \min_{x\in\mathbb{R}^m} \|b - Ax\|_2.$$

Applications:

Definition

Consider the possibly overdetermined system of equations

$$Ax = b \tag{1}$$

where $A \in \mathbb{R}^{n \times m}$ and a vector $b \in \mathbb{R}^n$, $n \geq m$. It may not have an exact solution if $n > m$ or A is square but singular. The **Least Squares Problem** associated with this system is about finding $x_0 \in \mathbb{R}^m$ such that

$$\|b - Ax_0\|_2 = \min_{x \in \mathbb{R}^m} \|b - Ax\|_2.$$

Applications:

- Polynomial curve fitting.

Definition

Consider the possibly overdetermined system of equations

$$Ax = b \tag{1}$$

where $A \in \mathbb{R}^{n \times m}$ and a vector $b \in \mathbb{R}^n$, $n \geq m$. It may not have an exact solution if $n > m$ or A is square but singular. The **Least Squares Problem** associated with this system is about finding $x_0 \in \mathbb{R}^m$ such that

$$\|b - Ax_0\|_2 = \min_{x \in \mathbb{R}^m} \|b - Ax\|_2.$$

Applications:

- Polynomial curve fitting.
- Making predictions from existing data.

Definition

Consider the possibly overdetermined system of equations

$$Ax = b \tag{1}$$

where $A \in \mathbb{R}^{n \times m}$ and a vector $b \in \mathbb{R}^n$, $n \geq m$. It may not have an exact solution if $n > m$ or A is square but singular. The **Least Squares Problem** associated with this system is about finding $x_0 \in \mathbb{R}^m$ such that

$$\|b - Ax_0\|_2 = \min_{x \in \mathbb{R}^m} \|b - Ax\|_2.$$

Applications:

- Polynomial curve fitting.
- Making predictions from existing data.
- Machine Learning.

Definition

Consider the possibly overdetermined system of equations

$$Ax = b \tag{1}$$

where $A \in \mathbb{R}^{n \times m}$ and a vector $b \in \mathbb{R}^n$, $n \geq m$. It may not have an exact solution if $n > m$ or A is square but singular. The **Least Squares Problem** associated with this system is about finding $x_0 \in \mathbb{R}^m$ such that

$$\|b - Ax_0\|_2 = \min_{x \in \mathbb{R}^m} \|b - Ax\|_2.$$

Applications:

- Polynomial curve fitting.
- Making predictions from existing data.
- Machine Learning.

Direct solution methods:

Definition

Consider the possibly overdetermined system of equations

$$Ax = b \tag{1}$$

where $A \in \mathbb{R}^{n \times m}$ and a vector $b \in \mathbb{R}^n$, $n \geq m$. It may not have an exact solution if $n > m$ or A is square but singular. The **Least Squares Problem** associated with this system is about finding $x_0 \in \mathbb{R}^m$ such that

$$\|b - Ax_0\|_2 = \min_{x \in \mathbb{R}^m} \|b - Ax\|_2.$$

Applications:

- Polynomial curve fitting.
- Making predictions from existing data.
- Machine Learning.

Direct solution methods:

- Normal Equations Method.

Definition

Consider the possibly overdetermined system of equations

$$Ax = b \tag{1}$$

where $A \in \mathbb{R}^{n \times m}$ and a vector $b \in \mathbb{R}^n$, $n \geq m$. It may not have an exact solution if $n > m$ or A is square but singular. The **Least Squares Problem** associated with this system is about finding $x_0 \in \mathbb{R}^m$ such that

$$\|b - Ax_0\|_2 = \min_{x \in \mathbb{R}^m} \|b - Ax\|_2.$$

Applications:

- Polynomial curve fitting.
- Making predictions from existing data.
- Machine Learning.

Direct solution methods:

- Normal Equations Method.
- QR Decomposition Method.

Definition

Consider the possibly overdetermined system of equations

$$Ax = b \tag{1}$$

where $A \in \mathbb{R}^{n\times m}$ and a vector $b \in \mathbb{R}^n$, $n \geq m$. It may not have an exact solution if $n > m$ or A is square but singular. The **Least Squares Problem** associated with this system is about finding $x_0 \in \mathbb{R}^m$ such that

$$\|b - Ax_0\|_2 = \min_{x\in\mathbb{R}^m} \|b - Ax\|_2.$$

Applications:

- Polynomial curve fitting.
- Making predictions from existing data.
- Machine Learning.

Direct solution methods:

- Normal Equations Method.
- QR Decomposition Method.
- Singular Value Decomposition Method.

Geometry of the Least Squares Problem

As $A \in \mathbb{R}^{n\times m}$, the linear map from $\mathbb{R}^m$ to $\mathbb{R}^n$ given by $x \mapsto Ax$, for all $x \in \mathbb{R}^m$ has range and null spaces

$$R(A) = \underbrace{\{Ax : x \in \mathbb{R}^m\}}_{\text{also called Col (A)}} \subset \mathbb{R}^n \text{ and } N(A) = \{x \in \mathbb{R}^m : Ax = 0\} \subset \mathbb{R}^m.$$

Geometry of the Least Squares Problem

As $A \in \mathbb{R}^{n \times m}$, the linear map from $\mathbb{R}^m$ to $\mathbb{R}^n$ given by $x \mapsto Ax$, for all $x \in \mathbb{R}^m$ has range and null spaces

$$R(A) = \underbrace{\{Ax : x \in \mathbb{R}^m\}}_{\text{also called Col (A)}} \subset \mathbb{R}^n \text{ and } N(A) = \{x \in \mathbb{R}^m : Ax = 0\} \subset \mathbb{R}^m.$$

The LSP is a two stage process involving

1. Find the best approximation to b from $R(A)$ i.e., find $y_0 \in R(A)$ such that $\|b - y_0\|_2 = \min_{y \in R(A)} \|b - y\|_2$.
2. Find $x_0 \in \mathbb{R}^m$ such that $Ax_0 = y_0$.

Geometry of the Least Squares Problem

As $A \in \mathbb{R}^{n\times m}$, the linear map from $\mathbb{R}^m$ to $\mathbb{R}^n$ given by $x \mapsto Ax$, for all $x \in \mathbb{R}^m$ has range and null spaces

$$R(A) = \underbrace{\{Ax : x \in \mathbb{R}^m\}}_{\text{also called Col (A)}} \subset \mathbb{R}^n \text{ and } N(A) = \{x \in \mathbb{R}^m : Ax = 0\} \subset \mathbb{R}^m.$$

The LSP is a two stage process involving

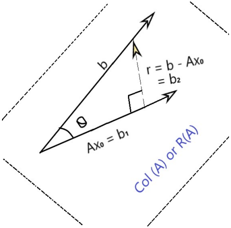
1. Find the best approximation to b from $R(A)$ i.e., find $y_0 \in R(A)$ such that $\|b - y_0\|_2 = \min_{y\in R(A)} \|b - y\|_2$.
2. Find $x_0 \in \mathbb{R}^m$ such that $Ax_0 = y_0$.

Since

$$\mathbb{R}^n = R(A) \oplus R(A)^\perp,$$

for $b \in R^n$, there exists unique $b_1 \in R(A)$, $b_2 \in R(A)^\perp$, such that $b = b_1 + b_2$.

Geometry of the Least Squares Problem



Therefore for all $x \in \mathbb{R}^m$, $b - Ax = \underbrace{b_1 - Ax}_{\in R(A)} + \underbrace{b_2}_{\in R(A)^\perp}$ and

$$\begin{aligned}
\|b - Ax\|_2^2 &= \langle b_1 - Ax + b_2, b_1 - Ax + b_2 \rangle \\
&= \|b_1 - Ax\|_2^2 + \|b_2\|_2^2 \qquad \text{(as } \langle b_1 - Ax, b_2 \rangle = 0 \text{ for all } x \in \mathbb{R}^m) \\
&\geq \|b_2\|_2^2 \\
&= \| \underbrace{b - Ax_0}_{:=r(=b_2)} \|_2^2
\end{aligned}$$

where $x_0 \in \mathbb{R}^m$ such that $Ax_0 = b_1$. Hence,

$$\min_{x \in \mathbb{R}^m} \|b - Ax\|_2^2 \geq \|b - Ax_0\|_2^2 \geq \min_{x \in \mathbb{R}^m} \|b - Ax\|_2^2 \Rightarrow \min_{x \in \mathbb{R}^m} \|b - Ax\|_2 = \|b - Ax_0\|_2.$$

Geometry of the Least Squares Problem

Hence $x_0 \in \mathbb{R}^m$ is a solution of the least squares problem and

$$r := b - Ax_0 = b_2 \in R(A)^\perp$$

is the residual vector.

Note that if P be the orthogonal projection of $\mathbb{R}^n$ onto $R(A)$, then $Pb = b_1$. Therefore in summary,

> *the orthogonal projection b_1 of b onto the column space of A is the nearest vector to b in the column space of A with respect to $\|\cdot\|_2$ and the solution of the LSP associated with the system is a column vector x_0 of scalars which produce the linear combination of the columns of A to form b_1.*

Geometry of the Least Squares Problem

Hence $x_0 \in \mathbb{R}^m$ is a solution of the least squares problem and

$$r := b - Ax_0 = b_2 \in R(A)^\perp$$

is the residual vector.

Note that if P be the orthogonal projection of $\mathbb{R}^n$ onto $R(A)$, then $Pb = b_1$. Therefore in summary,

> *the orthogonal projection b_1 of b onto the column space of A is the nearest vector to b in the column space of A with respect to $\|\cdot\|_2$ and the solution of the LSP associated with the system is a column vector x_0 of scalars which produce the linear combination of the columns of A to form b_1.*

Hence the solution of an LSP problem always exists!

Geometry of the Least Squares Problem

Hence $x_0 \in \mathbb{R}^m$ is a solution of the least squares problem and

$$r := b - Ax_0 = b_2 \in R(A)^\perp$$

is the residual vector.

Note that if P be the orthogonal projection of $\mathbb{R}^n$ onto $R(A)$, then $Pb = b_1$. Therefore in summary,

> *the orthogonal projection b_1 of b onto the column space of A is the nearest vector to b in the column space of A with respect to $\|\cdot\|_2$ and the solution of the LSP associated with the system is a column vector x_0 of scalars which produce the linear combination of the columns of A to form b_1.*

Hence the solution of an LSP problem always exists!

How many solutions are there?

Geometry of the Least Squares Problem

Hence $x_0 \in \mathbb{R}^m$ is a solution of the least squares problem and

$$r := b - Ax_0 = b_2 \in R(A)^{\perp}$$

is the residual vector.

Note that if P be the orthogonal projection of $\mathbb{R}^n$ onto $R(A)$, then $Pb = b_1$. Therefore in summary,

> *the orthogonal projection b_1 of b onto the column space of A is the nearest vector to b in the column space of A with respect to $\|\cdot\|_2$ and the solution of the LSP associated with the system is a column vector x_0 of scalars which produce the linear combination of the columns of A to form b_1.*

Hence the solution of an LSP problem always exists!

How many solutions are there?

The solution is unique if and only if $\operatorname{rank} A = m$.

Geometry of the Least Squares Problem

Hence $x_0 \in \mathbb{R}^m$ is a solution of the least squares problem and

$$r := b - Ax_0 = b_2 \in R(A)^\perp$$

is the residual vector.

Note that if P be the orthogonal projection of $\mathbb{R}^n$ onto $R(A)$, then $Pb = b_1$. Therefore in summary,

> *the orthogonal projection b_1 of b onto the column space of A is the nearest vector to b in the column space of A with respect to $\|\cdot\|_2$ and the solution of the LSP associated with the system is a column vector x_0 of scalars which produce the linear combination of the columns of A to form b_1.*

Hence the solution of an LSP problem always exists!

How many solutions are there?

The solution is unique if and only if $\operatorname{rank} A = m$.
If $\operatorname{rank} A < m$, there are infinitely many solutions!

Geometry of the Least Squares Problem

Hence $x_0 \in \mathbb{R}^m$ is a solution of the least squares problem and

$$r := b - Ax_0 = b_2 \in R(A)^\perp$$

is the residual vector.

Note that if P be the orthogonal projection of $\mathbb{R}^n$ onto $R(A)$, then $Pb = b_1$. Therefore in summary,

> *the orthogonal projection b_1 of b onto the column space of A is the nearest vector to b in the column space of A with respect to $\|\cdot\|_2$ and the solution of the LSP associated with the system is a column vector x_0 of scalars which produce the linear combination of the columns of A to form b_1.*

Hence the solution of an LSP problem always exists!

How many solutions are there?

The solution is unique if and only if $\operatorname{rank} A = m$.
If $\operatorname{rank} A < m$, there are infinitely many solutions!

Exercise: Prove the above statements!

Normal Equations Method

How to get x_0?

Since $R(A)^\perp = N(A^T)$, where $N(A^T)$ is the null space of the linear map $y \mapsto A^T y$ from $\mathbb{R}^n$ to $\mathbb{R}^m$, $A^T r = 0$.

Normal Equations Method

How to get x_0?

Since $R(A)^{\perp} = N(A^T)$, where $N(A^T)$ is the null space of the linear map $y \mapsto A^T y$ from $\mathbb{R}^n$ to $\mathbb{R}^m$, $A^T r = 0$. Hence x_0 is a solution of

$$A^T A x = A^T b \tag{2}$$

The systems of equations (2) are called the *Normal Equations* of the Least Squares Problem associated with $Ax = b$.

Normal Equations Method

How to get x_0?

Since $R(A)^{\perp} = N(A^T)$, where $N(A^T)$ is the null space of the linear map $y \mapsto A^T y$ from $\mathbb{R}^n$ to $\mathbb{R}^m$, $A^T r = 0$. Hence x_0 is a solution of

$$A^T A x = A^T b \tag{2}$$

The systems of equations (2) are called the *Normal Equations* of the Least Squares Problem associated with $Ax = b$.

Exercise: Let $A \in \mathbb{R}^{n \times m}$ and $b \in \mathbb{R}^n$, where $n \geq m$. Prove that the following statements are equivalent.

1. $\operatorname{rank} A = m$.
2. $A^T A$ is nonsingular.
3. $A^T A$ is positive definite.

Normal Equations Method

How to get x_0?

Since $R(A)^{\perp} = N(A^T)$, where $N(A^T)$ is the null space of the linear map $y \mapsto A^T y$ from $\mathbb{R}^n$ to $\mathbb{R}^m$, $A^T r = 0$. Hence x_0 is a solution of

$$A^T A x = A^T b \tag{2}$$

The systems of equations (2) are called the *Normal Equations* of the Least Squares Problem associated with $Ax = b$.

Exercise: Let $A \in \mathbb{R}^{n \times m}$ and $b \in \mathbb{R}^n$, where $n \geq m$. Prove that the following statements are equivalent.

1. $\mathrm{rank}\, A = m$.
2. $A^T A$ is nonsingular.
3. $A^T A$ is positive definite.

Pseudocode for solving the Least Squares Problem via Normal Equations when $\mathrm{rank}\, A = m$:

1. Form the Normal Equations. $(2nm^2 + O(nm)$ flops$)$
2. Solve them via the Cholesky Method. $(m^3/3 + O(m^2)$ flops$)$
3. Compute two norm squared of the residual vector. $(2nm + 3n$ flops$)$

Normal Equations Method

Since $\kappa_2(A^T A) = (\kappa_2(A))^2$, the Normal Equations can be very ill conditioned even if A is mildly ill conditioned.

Normal Equations Method

Since $\kappa_2(A^TA) = (\kappa_2(A))^2$, the Normal Equations can be very ill conditioned even if A is mildly ill conditioned.

The computation of A^TA is the most expensive step in the process and is also very prone to rounding error.

For example, for the full rank matrix

$$A = \begin{bmatrix} 1 & 1 \\ \epsilon & 0 \\ 0 & \epsilon \end{bmatrix}$$

where $\epsilon^2 < u$, $\mathrm{fl}(A^TA) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ is singular!

Therefore the method may not perform well even for mildly ill conditioned matrices.

QR Decomposition Method

Case 1: A is full rank, i.e., $\mathrm{rank}\, A = m$.

Let $A = Q \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$ be a QR decomposition of A and Q_1 be the isometry formed by the first m columns of Q.

QR Decomposition Method

Case 1: A is full rank, i.e., $\operatorname{rank} A = m$.

Let $A = Q\begin{bmatrix} R_1 \\ 0 \end{bmatrix}$ be a QR decomposition of A and Q_1 be the isometry formed by the first m columns of Q.

Then $Q_1 Q_1^T$ is the orthogonal projection onto $R(A)$ or $Col(A)$. Hence, $b_1 = Q_1 Q_1^T b$ and the solution x_0 of the LSP associated with $Ax = b$ satisfies

$$
\begin{aligned}
& Ax_0 = Q_1 Q_1^T b \\
\Rightarrow & Q\begin{bmatrix} R_1 \\ 0 \end{bmatrix} x_0 = Q[e_1 \ \cdots \ e_m][e_1 \ \cdots \ e_m]^T \underbrace{Q^T b}_{:=c} \\
\Rightarrow & \begin{bmatrix} R_1 \\ 0 \end{bmatrix} x_0 = \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix}_{n\times n} c \\
\Rightarrow & R_1 x_0 = c_1
\end{aligned}
$$

where c_1 is the vector of first m rows of c.

QR Decomposition Method

Case 1: A is full rank, i.e., $\operatorname{rank} A = m$.

Let $A = Q \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$ be a QR decomposition of A and Q_1 be the isometry formed by the first m columns of Q.

Then $Q_1 Q_1^T$ is the orthogonal projection onto $R(A)$ or $Col(A)$. Hence, $b_1 = Q_1 Q_1^T b$ and the solution x_0 of the LSP associated with $Ax = b$ satisfies

$$\begin{aligned}
&Ax_0 = Q_1 Q_1^T b \\
\Rightarrow &Q \begin{bmatrix} R_1 \\ 0 \end{bmatrix} x_0 = Q[e_1 \ \cdots \ e_m][e_1 \ \cdots \ e_m]^T \underbrace{Q^T b}_{:=c} \\
\Rightarrow &\begin{bmatrix} R_1 \\ 0 \end{bmatrix} x_0 = \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix}_{n \times n} c \\
\Rightarrow &R_1 x_0 = c_1
\end{aligned}$$

where c_1 is the vector of first m rows of c.

QR Decomposition Method

Case 2: $\operatorname{rank} A = r < m$.

QR Decomposition Method

Case 2: $\text{rank}\, A = r < m$.

Exercise: Let $AP = Q \begin{bmatrix} R_1 & R_2 \\ 0 & 0 \end{bmatrix}$ be a rank revealing QR decomposition of A where $R_1 \in \mathbb{R}^{r \times r}$, $R_2 \in \mathbb{R}^{r \times m-r}$, R_1 being upper triangular and nonsingular. Let $Q_1 = [q_1 \ \cdots \ q_r]$ be the isometry formed by the first r columns of Q. Then prove that

$$Col(A)(= R(A)) = \text{span}\,\{APe_1, \cdots, APe_r\} = \text{span}\,\{q_1, \cdots, q_r\} = Col(Q_1).$$

QR Decomposition Method

Case 2: $\text{rank}\, A = r < m$.

Exercise: Let $AP = Q \begin{bmatrix} R_1 & R_2 \\ 0 & 0 \end{bmatrix}$ be a rank revealing QR decomposition of A where $R_1 \in \mathbb{R}^{r\times r}$, $R_2 \in \mathbb{R}^{r\times m-r}$, R_1 being upper triangular and nonsingular. Let $Q_1 = [q_1 \ \cdots \ q_r]$ be the isometry formed by the first r columns of Q. Then prove that

$$Col(A)(= R(A)) = \text{span}\,\{APe_1, \cdots, APe_r\} = \text{span}\,\{q_1, \cdots, q_r\} = Col(Q_1).$$

Therefore, $Q_1 Q_1^T$ is the orthogonal projection onto $Col(A)$ and any solution x_0 of the LSP associated with $Ax = b$ satisfies

$$Ax_0 = Q_1 Q_1^T b \Rightarrow Q \begin{bmatrix} R_1 & R_2 \\ 0 & 0 \end{bmatrix} \underbrace{P^T x_0}_{:=y} = Q[e_1 \ \cdots \ e_r][e_1 \ \cdots \ e_r]^T \underbrace{Q^T b}_{:=c}$$

$$\Rightarrow \begin{bmatrix} R_1 & R_2 \\ 0 & 0 \end{bmatrix}_{n\times m} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{m\times 1} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{n\times n} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}_{n\times 1}$$

c_1, y_1 being the vectors of first r rows of c and y respectively and c_2, y_2 the vector formed by the last $m - r$ rows of y and last $n - r$ rows of c respectively

QR Decomposition Method

Then, $R_1 y_1 + R_2 y_2 = c_1 \Rightarrow R_1 y_1 = c_1 - R_2 y_2$.

QR Decomposition Method

Then, $R_1 y_1 + R_2 y_2 = c_1 \Rightarrow R_1 y_1 = c_1 - R_2 y_2$.

Choose y_2 arbitrarily and y_1 as the solution of the above $r \times r$ upper triangular system.

QR Decomposition Method

Then, $R_1 y_1 + R_2 y_2 = c_1 \Rightarrow R_1 y_1 = c_1 - R_2 y_2$.

Choose y_2 arbitrarily and y_1 as the solution of the above $r \times r$ upper triangular system.

For ease of computation choose $y_2 = 0$. Then one solution $x_0 \in \mathbb{R}^m$ satisfying

$$\|b - Ax_0\|_2^2 = \min_{x \in \mathbb{R}^m} \|b - Ax\|_2^2$$

is given by

$$x_0 = P \begin{bmatrix} y_1 \\ 0 \end{bmatrix}$$

where $R_1 y_1 = c_1$.

QR Decomposition Method

Then, $R_1 y_1 + R_2 y_2 = c_1 \Rightarrow R_1 y_1 = c_1 - R_2 y_2$.

Choose y_2 arbitrarily and y_1 as the solution of the above $r \times r$ upper triangular system.

For ease of computation choose $y_2 = 0$. Then one solution $x_0 \in \mathbb{R}^m$ satisfying

$$\|b - Ax_0\|_2^2 = \min_{x \in \mathbb{R}^m} \|b - Ax\|_2^2$$

is given by

$$x_0 = P \begin{bmatrix} y_1 \\ 0 \end{bmatrix}$$

where $R_1 y_1 = c_1$.

Exercise: If $r = b - Ax_0$ prove that $\|r\|_2 = \|c_2\|_2$.

QR Decomposition Method

Pseudocode for solving the LSP associated with $Ax = b$ via QR decomposition method:

1. Find $R = \begin{bmatrix} R_1 & R_2 \\ 0 & 0 \end{bmatrix}$ of a rank revealing QR decomposition of A and suppose $Q_1, \ldots, Q_r$ are the reflectors and $P_1, \ldots P_r$ are the permutations required in the process. Here $r \le m$.
   (costs $2nm^2 - \frac{2}{3}m^3 + O(nm) + O(m^2)$ flops if $r = m$.)

QR Decomposition Method

Pseudocode for solving the LSP associated with $Ax = b$ via QR decomposition method:

1. Find $R = \begin{bmatrix} R_1 & R_2 \\ 0 & 0 \end{bmatrix}$ of a rank revealing QR decomposition of A and suppose $Q_1, \ldots, Q_r$ are the reflectors and $P_1, \ldots P_r$ are the permutations required in the process. Here $r \leq m$.
   (costs $2nm^2 - \frac{2}{3}m^3 + O(nm) + O(m^2)$ flops if $r = m$.)
2. Compute $c = Q_r \cdots Q_1 b$ and extract vectors c_1, c_2 from its first entries.
   (costs $O(nm)$ flops if $r = m$.)

QR Decomposition Method

Pseudocode for solving the LSP associated with $Ax = b$ via QR decomposition method:

1. Find $R = \begin{bmatrix} R_1 & R_2 \\ 0 & 0 \end{bmatrix}$ of a rank revealing QR decomposition of A and suppose $Q_1, \ldots, Q_r$ are the reflectors and $P_1, \ldots P_r$ are the permutations required in the process. Here $r \leq m$.
   (costs $2nm^2 - \frac{2}{3}m^3 + O(nm) + O(m^2)$ flops if $r = m$.)
2. Compute $c = Q_r \cdots Q_1 b$ and extract vectors c_1, c_2 from its first entries.
   (costs $O(nm)$ flops if $r = m$.)
3. Find $\|r\|_2^2 = \|c_2\|_2^2$.
   (costs $O(nm)$ flops if $r = m$.)

QR Decomposition Method

Pseudocode for solving the LSP associated with $Ax = b$ via QR decomposition method:

1. Find $R = \begin{bmatrix} R_1 & R_2 \\ 0 & 0 \end{bmatrix}$ of a rank revealing QR decomposition of A and suppose $Q_1, \ldots, Q_r$ are the reflectors and $P_1, \ldots P_r$ are the permutations required in the process. Here $r \leq m$.
   (costs $2nm^2 - \frac{2}{3}m^3 + O(nm) + O(m^2)$ flops if $r = m$.)
2. Compute $c = Q_r \cdots Q_1 b$ and extract vectors c_1, c_2 from its first entries.
   (costs $O(nm)$ flops if $r = m$.)
3. Find $\|r\|_2^2 = \|c_2\|_2^2$.
   (costs $O(nm)$ flops if $r = m$.)
4. Solve $R_1 y = c_1$.
   (costs $O(m^2)$ flops if $r = m$.)

QR Decomposition Method

Pseudocode for solving the LSP associated with $Ax = b$ via QR decomposition method:

1. Find $R = \begin{bmatrix} R_1 & R_2 \\ 0 & 0 \end{bmatrix}$ of a rank revealing QR decomposition of A and suppose $Q_1, \ldots, Q_r$ are the reflectors and $P_1, \ldots P_r$ are the permutations required in the process. Here $r \leq m$.
   (costs $2nm^2 - \frac{2}{3}m^3 + O(nm) + O(m^2)$ flops if $r = m$.)
2. Compute $c = Q_r \cdots Q_1 b$ and extract vectors c_1, c_2 from its first entries.
   (costs $O(nm)$ flops if $r = m$.)
3. Find $\|r\|_2^2 = \|c_2\|_2^2$.
   (costs $O(nm)$ flops if $r = m$.)
4. Solve $R_1 y = c_1$.
   (costs $O(m^2)$ flops if $r = m$.)
5. Set $x_0 = P_1 \cdots P_r \begin{bmatrix} y \\ 0 \end{bmatrix}_{m \times 1}$.
   (Here the zeroes are necessary only if $r < m$.)