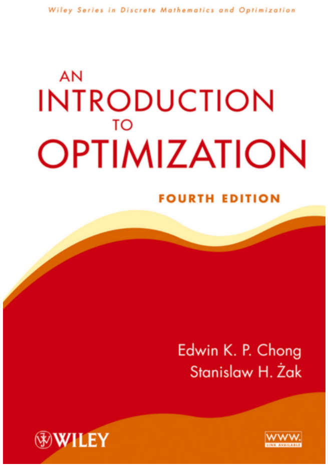


Introduction to Optimization

K. R. Sahasranand

Data Science

sahasranand@iitpkd.ac.in

Recall

Suppose that we are given m equations in n unknowns of the form

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

In matrix form,

$$Ax = b.$$

◦ The system $Ax = b$ has a solution iff rank A = rank $[A,\ b]$.

Tall matrices

≈ more rows than columns

Consider a system of linear equations

$$Ax = b$$

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, m \geq n$, and rank $A = n$.

$$\begin{bmatrix} 0 & 8 & 44 \\ 1 & 13 & 33 \\ 1 & 21 & 77 \\ 2 & 34 & 10 \\ 3 & 55 & 87 \\ 5 & 89 & 97 \end{bmatrix}$$

Tall matrices

≈ more rows than columns

Consider a system of linear equations

$$Ax = b$$

where $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, m \geq n$, and rank $A = n$.

$$\begin{bmatrix} 0 & 8 & 44 \\ 1 & 13 & 33 \\ 1 & 21 & 77 \\ 2 & 34 & 10 \\ 3 & 55 & 87 \\ 5 & 89 & 97 \end{bmatrix}$$

- $\mathrm{Col}(A) \neq \mathbb{R}^m$

Tall matrices

≈ more rows than columns

Consider a system of linear equations

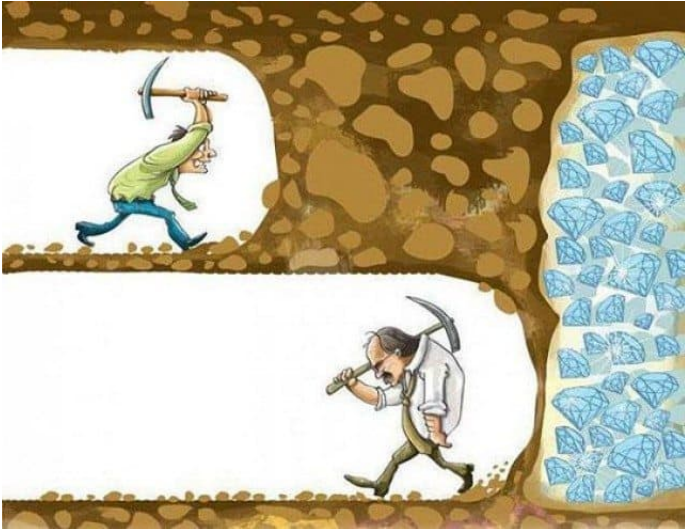
$$Ax = b$$

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, m \geq n$, and rank $A = n$.

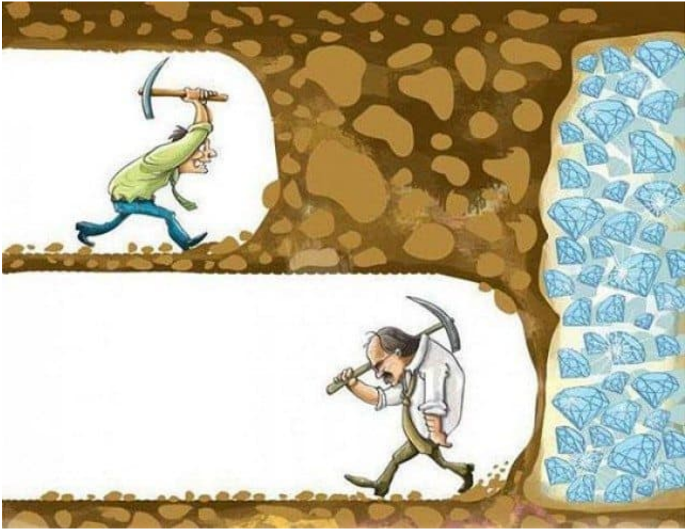
$$\begin{bmatrix} 0 & 8 & 44 \\ 1 & 13 & 33 \\ 1 & 21 & 77 \\ 2 & 34 & 10 \\ 3 & 55 & 87 \\ 5 & 89 & 97 \end{bmatrix}$$

- $\mathrm{Col}(A) \neq \mathbb{R}^m$
- What if $b \notin \mathrm{Col}(A)$?

No solution?



No solution?



~~No solution~~ Least squares solution

Least squares solution

$A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, m \geq n$, and rank $A = n$

Find the vector x that minimizes $\|Ax - b\|^2$

Least squares solution

$A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, m \geq n$, and rank $A = n$

Find the vector x that minimizes $\|Ax - b\|^2$

$\equiv$ Find x^* such that for all $x \in \mathbb{R}^n$,

$$\|Ax - b\|^2 \geq \|Ax^* - b\|^2$$

Least squares solution

$A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, m \geq n$, and rank $A = n$

Find the vector x that minimizes $\|Ax - b\|^2$

$\equiv$ Find x^* such that for all $x \in \mathbb{R}^n$,

$$\|Ax - b\|^2 \geq \|Ax^* - b\|^2$$

x^* is the least squares solution to $Ax = b$.

Helper lemma

for finding the least squares solution

Lemma – Let $A \in \mathbb{R}^{m \times n}, m \geq n$. Then,

$$\text{rank } A = n \iff \text{rank } A^\top A = n$$

Helper lemma

for finding the least squares solution

Lemma – Let $A \in \mathbb{R}^{m \times n}, m \geq n$. Then,

$$\text{rank } A = n \iff \text{rank } A^\top A = n$$

Proof sketch –

($\Longrightarrow$) Suppose rank $A = n$. Show that $\mathcal{N}(A^\top A) = \{\mathbf{0}\}$.

Let $x \in \mathcal{N}(A^\top A)$. Then,

$$\|Ax\|^2 = x^\top A^\top Ax$$

Helper lemma

for finding the least squares solution

Lemma – Let $A \in \mathbb{R}^{m \times n}, m \geq n$. Then,

$$\text{rank } A = n \iff \text{rank } A^\top A = n$$

Proof sketch –

$(\Longrightarrow)$ Suppose rank $A = n$. Show that $\mathcal{N}(A^\top A) = \{\mathbf{0}\}$.

Let $x \in \mathcal{N}(A^\top A)$. Then,

$$\|Ax\|^2 = x^\top A^\top Ax = \mathbf{0} \quad \Longrightarrow \quad Ax = \mathbf{0}.$$

Helper lemma

for finding the least squares solution

Lemma – Let $A \in \mathbb{R}^{m \times n}, m \geq n$. Then,

$$\text{rank } A = n \iff \text{rank } A^\top A = n$$

Proof sketch –

($\Longrightarrow$) Suppose rank $A = n$. Show that $\mathcal{N}(A^\top A) = \{\mathbf{0}\}$.

Let $x \in \mathcal{N}(A^\top A)$. Then,

$$\|Ax\|^2 = x^\top A^\top Ax = \mathbf{0} \quad \Longrightarrow \quad Ax = \mathbf{0}.$$

($\Longleftarrow$) Suppose rank $A^\top A = n$. Show that $\mathcal{N}(A) = \{\mathbf{0}\}$.

Helper lemma

for finding the least squares solution

Lemma – Let $A \in \mathbb{R}^{m \times n}, m \geq n$. Then,

$$\text{rank } A = n \iff \text{rank } A^\top A = n$$

Proof sketch –

$(\Longrightarrow)$ Suppose rank $A = n$. Show that $\mathcal{N}(A^\top A) = \{\mathbf{0}\}$.

Let $x \in \mathcal{N}(A^\top A)$. Then,

$$\|Ax\|^2 = x^\top A^\top Ax = \mathbf{0} \quad \Longrightarrow \quad Ax = \mathbf{0}.$$

$(\Longleftarrow)$ Suppose rank $A^\top A = n$. Show that $\mathcal{N}(A) = \{\mathbf{0}\}$.

Let $x \in \mathcal{N}(A)$. Then,

$$Ax = \mathbf{0}$$

Helper lemma

for finding the least squares solution

Lemma – Let $A \in \mathbb{R}^{m\times n}, m \geq n$. Then,

$$\text{rank } A = n \iff \text{rank } A^\top A = n$$

Proof sketch –

($\Longrightarrow$) Suppose rank $A = n$. Show that $\mathcal{N}(A^\top A) = \{\mathbf{0}\}$.

Let $x \in \mathcal{N}(A^\top A)$. Then,

$$\|Ax\|^2 = x^\top A^\top Ax = \mathbf{0} \quad \Longrightarrow \quad Ax = \mathbf{0}.$$

($\Longleftarrow$) Suppose rank $A^\top A = n$. Show that $\mathcal{N}(A) = \{\mathbf{0}\}$.

Let $x \in \mathcal{N}(A)$. Then,

$$Ax = \mathbf{0} \quad \Longrightarrow \quad A^\top Ax = \mathbf{0}.$$

Least squares solution

$A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, m \geq n$, and rank $A = n$

Find the vector x that minimizes $\|Ax - b\|^2$.

Least squares solution

$A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, m \geq n$, and rank $A = n$

Find the vector x that minimizes $\|Ax - b\|^2$.

Theorem – The unique vector x^* that minimizes $\|Ax - b\|^2$ is given by the solution to the equation

$$A^\top A x = A^\top b.$$

That is,

$$x^* = (A^\top A)^{-1} A^\top b.$$

Least squares solution

$A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, m \geq n$, and rank $A = n$

Find the vector x that minimizes $\|Ax - b\|^2$.

Theorem – The unique vector x^* that minimizes $\|Ax - b\|^2$ is given by the solution to the equation

$$A^\top A x = A^\top b.$$

That is,

$$x^* = (A^\top A)^{-1} A^\top b.$$

Proof sketch – First show that

$$\begin{aligned} \|Ax - b\|^2 &= \|Ax - Ax^* + Ax^* - b\|^2 \\ &= \cdots \\ &= \|A(x - x^*)\|^2 + \|Ax^* - b\|^2. \end{aligned}$$

Least squares solution

$A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, m \geq n$, and rank $A = n$

Find the vector x that minimizes $\|Ax - b\|^2$.

Theorem – The unique vector x^* that minimizes $\|Ax - b\|^2$ is given by the solution to the equation

$$A^\top A x = A^\top b.$$

That is,

$$x^* = (A^\top A)^{-1} A^\top b.$$

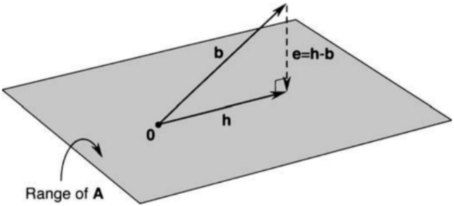
Proof sketch – First show that

$$\begin{aligned} \|Ax - b\|^2 &= \|Ax - Ax^* + Ax^* - b\|^2 \\ &= \cdots \\ &= \|A(x - x^*)\|^2 + \|Ax^* - b\|^2. \end{aligned}$$

If $x \neq x^*$, then $\|A(x - x^*)\|^2 > 0$ since rank $A = n$.

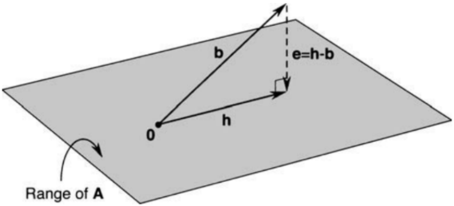
(explain on board)

Geometric picture of the least squares solution



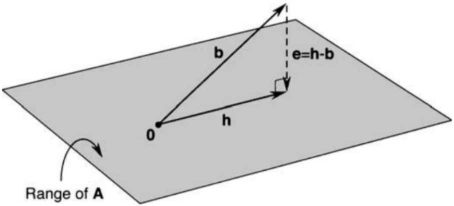
- Look for the vector $h \in \mathrm{Col}(A)$ closest to b

Geometric picture of the least squares solution



- Look for the vector $h \in \mathrm{Col}(A)$ closest to b
- Then, every column of A is orthogonal to the "error" $e = h - b$:

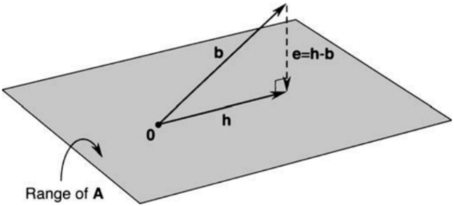
Geometric picture of the least squares solution



- Look for the vector $h \in \mathrm{Col}(A)$ closest to b
- Then, every column of A is orthogonal to the "error" $e = h - b$:

$$a_i^\top e = 0 \qquad ; \qquad a_i \sim i\text{th column of } A$$

Geometric picture of the least squares solution



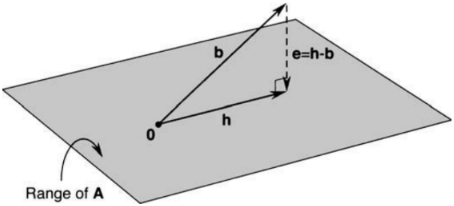
- Look for the vector $h \in \mathrm{Col}(A)$ closest to b

- Then, every column of A is orthogonal to the "error" $e = h - b$:

$$a_i^\top e = 0 \qquad ; \qquad a_i \sim i\text{th column of } A$$

  In matrix form, $A^\top e = \mathbf{0}$.

Geometric picture of the least squares solution



- Look for the vector $h \in \mathrm{Col}(A)$ closest to b
- Then, every column of A is orthogonal to the "error" $e = h - b$:

  $$a_i^\top e = 0 \qquad ; \qquad a_i \sim i\text{th column of } A$$

  In matrix form, $A^\top e = \mathbf{0}$.

- Since $h \in \mathrm{Col}(A)$, we can write

  $$h = Ax^* \qquad \text{for some } x^*.$$

(continue on board)

A 2D example

(show on board)

Let

$$A = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

A 2D example

(show on board)

Let

$$A = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Example 1:

$$b = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad\qquad ;$$

A 2D example

(show on board)

Let

$$A = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Example 1:

$$b = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad ; \qquad b \notin \mathrm{Col}(A)$$

A 2D example

(show on board)

Let

$$A = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Example 1:

$$b = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad ; \qquad b \notin \mathrm{Col}(A)$$

Example 2:

$$b = \begin{bmatrix} 3 \\ 3/2 \end{bmatrix} \qquad ;$$

A 2D example

(show on board)

Let

$$A = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Example 1:

$$b = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad ; \qquad b \notin \mathrm{Col}(A)$$

Example 2:

$$b = \begin{bmatrix} 3 \\ 3/2 \end{bmatrix} \qquad ; \qquad b \in \mathrm{Col}(A)$$

Line fitting

We model test scores based on hours studied:

#hours studied	Test score
1	50
2	60
3	65
4	70

(plot on board)

Line fitting

We model test scores based on hours studied:

#hours studied	Test score
1	50
2	60
3	65
4	70

(plot on board)

Assuming $Ax = b$, this gives us:

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 50 \\ 60 \\ 65 \\ 70 \end{bmatrix} \qquad \text{with } x = \begin{bmatrix} m \\ c \end{bmatrix}$$

Least squares solution

We solve using the least squares formula:

$$x^* = (A^\top A)^{-1} A^\top b$$

Step-by-step:

$$A^\top A = \begin{bmatrix} 30 & 10 \\ 10 & 4 \end{bmatrix}, \quad A^\top b = \begin{bmatrix} 645 \\ 245 \end{bmatrix}$$

Least squares solution

We solve using the least squares formula:

$$x^* = (A^\top A)^{-1} A^\top b$$

Step-by-step:

$$A^\top A = \begin{bmatrix} 30 & 10 \\ 10 & 4 \end{bmatrix}, \quad A^\top b = \begin{bmatrix} 645 \\ 245 \end{bmatrix}$$

Inverse of $A^\top A$:

$$(A^\top A)^{-1} = \frac{1}{20} \begin{bmatrix} 4 & -10 \\ -10 & 30 \end{bmatrix}$$

Solving for $x = \begin{bmatrix} m \\ c \end{bmatrix}$

Now compute:

$$\begin{aligned} x^* &= (A^\top A)^{-1} A^\top b \\ &= \frac{1}{20} \begin{bmatrix} 4 & -10 \\ -10 & 30 \end{bmatrix} \begin{bmatrix} 645 \\ 245 \end{bmatrix} \end{aligned}$$

Solving for $x = \begin{bmatrix} m \\ c \end{bmatrix}$

Now compute:

$$\begin{aligned} x^* &= (A^\top A)^{-1} A^\top b \\ &= \frac{1}{20} \begin{bmatrix} 4 & -10 \\ -10 & 30 \end{bmatrix} \begin{bmatrix} 645 \\ 245 \end{bmatrix} \\ &= \begin{bmatrix} 6.5 \\ 45 \end{bmatrix} \end{aligned}$$

Solving for $x = \begin{bmatrix} m \\ c \end{bmatrix}$

Now compute:

$$\begin{aligned} x^* &= (A^\top A)^{-1} A^\top b \\ &= \frac{1}{20} \begin{bmatrix} 4 & -10 \\ -10 & 30 \end{bmatrix} \begin{bmatrix} 645 \\ 245 \end{bmatrix} \\ &= \begin{bmatrix} 6.5 \\ 45 \end{bmatrix} \end{aligned}$$

So, the best fit line is:

$$\boxed{y = 6.5t + 45}$$

Interpretation

predict test scores for "any" number of study hours

- $m = 6.5$: every hour of study $\Longrightarrow$ score of 6.5 points.
- $c = 45$: a student who doesn't study would score 45.

Interpretation

predict test scores for "any" number of study hours

- $m = 6.5$: every hour of study $\Longrightarrow$ score of 6.5 points.
- $c = 45$: a student who doesn't study would score 45.

- The prediction doesn't "match" the measurements!
  - for $t = 1$, $y = 51.5 \neq 50$, and so on.

Interpretation

predict test scores for "any" number of study hours

- $m = 6.5$: every hour of study $\implies$ score of 6.5 points.
- $c = 45$: a student who doesn't study would score 45.

- The prediction doesn't "match" the measurements!
  - for $t = 1$, $y = 51.5 \neq 50$, and so on.

- Suppose I get one more measurement:

#hours studied	Test score
1	50
2	60
3	65
4	70
5	92

Do the whole calculation from scratch?

Recursive least squares

Update the least squares solution x^* to accommodate new data point(s).

#hours studied	Test score
1	50
2	60
3	65
4	70

Recursive least squares

Update the least squares solution x^* to accommodate new data point(s).

#hours studied	Test score
1	50
2	60
3	65
4	70

– The solution to the least squares problem with A_0 and $b^{(0)}$ is given by

$$x^{(0)} = {G_0}^{-1} A_0^\top b^{(0)} \qquad ; \qquad G_0 = A_0^\top A_0$$

Recursive least squares

Update the least squares solution x^* to accommodate new data point(s).

#hours studied	Test score
1	50
2	60
3	65
4	70
5	92
6	88

– The solution to the least squares problem with A_0 and $b^{(0)}$ is given by

$$x^{(0)} = G_0{}^{-1} A_0^\top b^{(0)} \qquad ; \qquad G_0 = A_0^\top A_0$$

– Suppose we are given more data: A_1 and $b^{(1)}$.

Recursive least squares

General form

We seek to minimize

$$\left\| \begin{bmatrix} A_0 \\ A_1 \end{bmatrix} x - \begin{bmatrix} b^{(0)} \\ b^{(1)} \end{bmatrix} \right\|^2 .$$

Recursive least squares

General form

We seek to minimize

$$\left\| \begin{bmatrix} A_0 \\ A_1 \end{bmatrix} x - \begin{bmatrix} b^{(0)} \\ b^{(1)} \end{bmatrix} \right\|^2 .$$

If you think of this as $\|Bx - y\|^2$, the solution is given by $(B^\top B)^{-1} B^\top y$;

Recursive least squares

General form

We seek to minimize

$$\left\| \begin{bmatrix} A_0 \\ A_1 \end{bmatrix} x - \begin{bmatrix} b^{(0)} \\ b^{(1)} \end{bmatrix} \right\|^2 .$$

If you think of this as $\|Bx - y\|^2$, the solution is given by $(B^\top B)^{-1} B^\top y$;

that is

$$x^{(1)} = G_1^{-1} \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}^\top \begin{bmatrix} b^{(0)} \\ b^{(1)} \end{bmatrix}$$

where

$$G_1 = \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}^\top \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}$$

Recursive least squares

General form

We seek to minimize

$$\left\| \begin{bmatrix} A_0 \\ A_1 \end{bmatrix} x - \begin{bmatrix} b^{(0)} \\ b^{(1)} \end{bmatrix} \right\|^2 .$$

If you think of this as $\|Bx - y\|^2$, the solution is given by $(B^\top B)^{-1} B^\top y$;

that is

$$x^{(1)} = G_1^{-1} \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}^\top \begin{bmatrix} b^{(0)} \\ b^{(1)} \end{bmatrix}$$

where

$$G_1 = \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}^\top \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}$$

We want to write $x^{(1)}$ in terms of **old estimate**: $x^{(0)}$ and **new data**: $A_1, b^{(1)}$;
(possibly use G_0 also)

Recursive least squares

Update equations

$$x^{(1)} = G_1^{-1} \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}^\top \begin{bmatrix} b^{(0)} \\ b^{(1)} \end{bmatrix} \qquad ; \qquad G_1 = \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}^\top \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}$$

$$G_1 = \begin{bmatrix} A_0^\top & A_1^\top \end{bmatrix} \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}$$

Recursive least squares

Update equations

$$x^{(1)} = G_1^{-1} \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}^\top \begin{bmatrix} b^{(0)} \\ b^{(1)} \end{bmatrix} \qquad ; \qquad G_1 = \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}^\top \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}$$

$$\begin{aligned} G_1 &= \begin{bmatrix} A_0^\top & A_1^\top \end{bmatrix} \begin{bmatrix} A_0 \\ A_1 \end{bmatrix} \\ &= A_0^\top A_0 + A_1^\top A_1 \end{aligned}$$

Recursive least squares

Update equations

$$x^{(1)} = G_1^{-1} \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}^\top \begin{bmatrix} b^{(0)} \\ b^{(1)} \end{bmatrix} \qquad ; \qquad G_1 = \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}^\top \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}$$

$$\begin{aligned} G_1 &= \begin{bmatrix} A_0^\top & A_1^\top \end{bmatrix} \begin{bmatrix} A_0 \\ A_1 \end{bmatrix} \\ &= A_0^\top A_0 + A_1^\top A_1 \\ &= G_0 + A_1^\top A_1 \end{aligned}$$

Recursive least squares

Update equations

$$x^{(1)} = G_1^{-1} \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}^\top \begin{bmatrix} b^{(0)} \\ b^{(1)} \end{bmatrix} \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

$$\begin{aligned} G_1 &= \begin{bmatrix} A_0^\top & A_1^\top \end{bmatrix} \begin{bmatrix} A_0 \\ A_1 \end{bmatrix} \\ &= A_0^\top A_0 + A_1^\top A_1 \\ &= G_0 + A_1^\top A_1 \end{aligned}$$

Recursive least squares

Update equations

$$x^{(1)} = G_1^{-1} \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}^\top \begin{bmatrix} b^{(0)} \\ b^{(1)} \end{bmatrix} \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

$$\begin{bmatrix} A_0 \\ A_1 \end{bmatrix}^\top \begin{bmatrix} b^{(0)} \\ b^{(1)} \end{bmatrix} = \begin{bmatrix} A_0^\top & A_1^\top \end{bmatrix} \begin{bmatrix} b^{(0)} \\ b^{(1)} \end{bmatrix}$$

Recursive least squares

Update equations

$$x^{(1)} = G_1^{-1} \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}^\top \begin{bmatrix} b^{(0)} \\ b^{(1)} \end{bmatrix} \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

$$\begin{bmatrix} A_0 \\ A_1 \end{bmatrix}^\top \begin{bmatrix} b^{(0)} \\ b^{(1)} \end{bmatrix} = \begin{bmatrix} A_0^\top & A_1^\top \end{bmatrix} \begin{bmatrix} b^{(0)} \\ b^{(1)} \end{bmatrix}$$

$$= A_0^\top b^{(0)} + A_1^\top b^{(1)}.$$

Recursive least squares

Update equations

$$x^{(1)} = G_1^{-1}\left[A_0^\top b^{(0)} + A_1^\top b^{(1)}\right] \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

Recursive least squares

Update equations

$$x^{(1)} = G_1^{-1}\left[A_0^\top b^{(0)} + A_1^\top b^{(1)}\right] \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

Recursive least squares

Update equations

$$x^{(1)} = G_1^{-1}\left[A_0^\top b^{(0)} + A_1^\top b^{(1)}\right] \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

$$A_0^\top b^{(0)} = \left(G_0 G_0^{-1}\right) A_0^\top b^{(0)}$$

Recursive least squares

Update equations

$$x^{(1)} = G_1^{-1}\left[A_0^\top b^{(0)} + A_1^\top b^{(1)}\right] \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

$$\begin{aligned} A_0^\top b^{(0)} &= \left(G_0 G_0^{-1}\right) A_0^\top b^{(0)} \\ &= G_0 G_0^{-1} A_0^\top b^{(0)} \end{aligned}$$

Recursive least squares

Update equations

$$x^{(1)} = G_1^{-1}\left[A_0^\top b^{(0)} + A_1^\top b^{(1)}\right] \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

$$\begin{aligned} A_0^\top b^{(0)} &= \left(G_0 G_0^{-1}\right) A_0^\top b^{(0)} \\ &= G_0 G_0^{-1} A_0^\top b^{(0)} \\ &= G_0 x^{(0)} \end{aligned}$$

Recursive least squares

Update equations

$$x^{(1)} = G_1^{-1}\left[A_0^\top b^{(0)} + A_1^\top b^{(1)}\right] \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

$$\begin{aligned}
A_0^\top b^{(0)} &= \left(G_0 G_0^{-1}\right) A_0^\top b^{(0)} \\
&= G_0 G_0^{-1} A_0^\top b^{(0)} \\
&= G_0 x^{(0)} \\
&= \left(G_1 - A_1^\top A_1\right) x^{(0)}
\end{aligned}$$

Recursive least squares

Update equations

$$x^{(1)} = G_1^{-1}\left[A_0^\top b^{(0)} + A_1^\top b^{(1)}\right] \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

$$\begin{aligned}
A_0^\top b^{(0)} &= \left(G_0 G_0^{-1}\right) A_0^\top b^{(0)} \\
&= G_0 G_0^{-1} A_0^\top b^{(0)} \\
&= G_0 x^{(0)} \\
&= \left(G_1 - A_1^\top A_1\right) x^{(0)} \\
&= G_1 x^{(0)} - A_1^\top A_1 x^{(0)}.
\end{aligned}$$

Recursive least squares

Update equations

$$x^{(1)} = G_1^{-1}\left[G_1 x^{(0)} - A_1^\top A_1 x^{(0)} + A_1^\top b^{(1)}\right] \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

$$\begin{aligned}
A_0^\top b^{(0)} &= \left(G_0 G_0^{-1}\right) A_0^\top b^{(0)} \\
&= G_0 G_0^{-1} A_0^\top b^{(0)} \\
&= G_0 x^{(0)} \\
&= \left(G_1 - A_1^\top A_1\right) x^{(0)} \\
&= G_1 x^{(0)} - A_1^\top A_1 x^{(0)}.
\end{aligned}$$

Recursive least squares

Update equations

$$x^{(1)} = G_1^{-1}\Big[G_1 x^{(0)} - A_1^\top A_1 x^{(0)} + A_1^\top b^{(1)}\Big] \quad ; \quad G_1 = G_0 + A_1^\top A_1$$

$$x^{(1)} = G_1^{-1}\Big[G_1 x^{(0)} - A_1^\top A_1 x^{(0)} + A_1^\top b^{(1)}\Big]$$

Recursive least squares

Update equations

$$x^{(1)} = G_1^{-1}\left[G_1 x^{(0)} - A_1^\top A_1 x^{(0)} + A_1^\top b^{(1)}\right] \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

$$\begin{aligned} x^{(1)} &= G_1^{-1}\left[G_1 x^{(0)} - A_1^\top A_1 x^{(0)} + A_1^\top b^{(1)}\right] \\ &= x^{(0)} + G_1^{-1} A_1^\top \left[b^{(1)} - A_1 x^{(0)}\right] \end{aligned}$$

Recursive least squares

Update equations

$$x^{(1)} = G_1^{-1}\Big[G_1 x^{(0)} - A_1^\top A_1 x^{(0)} + A_1^\top b^{(1)}\Big] \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

$$\begin{aligned} x^{(1)} &= G_1^{-1}\Big[G_1 x^{(0)} - A_1^\top A_1 x^{(0)} + A_1^\top b^{(1)}\Big] \\ &= x^{(0)} + G_1{}^{-1} A_1^\top \Big[b^{(1)} - A_1 x^{(0)}\Big] \end{aligned}$$

Recursive least squares

Update equations

$$x^{(1)} = x^{(0)} + G_1{}^{-1} A_1^\top \left[b^{(1)} - A_1 x^{(0)} \right] \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

Recursive least squares

Update equations

$$x^{(1)} = x^{(0)} + {G_1}^{-1} A_1^\top \left[b^{(1)} - A_1 x^{(0)} \right] \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

Is there a way to compute G_1^{-1} from (the already computed) G_0^{-1}?

Recursive least squares

Update equations

$$x^{(1)} = x^{(0)} + {G_1}^{-1} A_1^\top \left[b^{(1)} - A_1 x^{(0)} \right] \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

Is there a way to compute G_1^{-1} from (the already computed) G_0^{-1}?

Sherman-Morrison-Woodbury formula:

$$(A + UV)^{-1} = A^{-1} - \left(A^{-1}U\right)\left(I + VA^{-1}U\right)^{-1}\left(VA^{-1}\right)$$

Recursive least squares

Update equations

$$x^{(1)} = x^{(0)} + {G_1}^{-1} A_1^\top \left[b^{(1)} - A_1 x^{(0)} \right] \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

Is there a way to compute G_1^{-1} from (the already computed) G_0^{-1}?

Sherman-Morrison-Woodbury formula: (verify by substitution)

$$(A + UV)^{-1} = A^{-1} - \left(A^{-1}U\right)\left(I + VA^{-1}U\right)^{-1}\left(VA^{-1}\right)$$

Recursive least squares

Update equations

$$x^{(1)} = x^{(0)} + {G_1}^{-1} A_1^\top \left[b^{(1)} - A_1 x^{(0)} \right] \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

Is there a way to compute G_1^{-1} from (the already computed) G_0^{-1}?

Sherman-Morrison-Woodbury formula: (verify by substitution)

$$(A + UV)^{-1} = A^{-1} - \left(A^{-1} U\right) \left(I + V A^{-1} U\right)^{-1} \left(V A^{-1}\right)$$

$$G_1^{-1} = (G_0 + A_1^\top A_1)^{-1}$$

Recursive least squares

Update equations

$$x^{(1)} = x^{(0)} + {G_1}^{-1} A_1^\top \left[b^{(1)} - A_1 x^{(0)} \right] \qquad ; \qquad G_1 = G_0 + A_1^\top A_1$$

Is there a way to compute G_1^{-1} from (the already computed) G_0^{-1}?

Sherman-Morrison-Woodbury formula: (verify by substitution)

$$(A + UV)^{-1} = A^{-1} - \left(A^{-1}U\right)\left(I + VA^{-1}U\right)^{-1}\left(VA^{-1}\right)$$

$$\begin{aligned} G_1^{-1} &= (G_0 + A_1^\top A_1)^{-1} \\ &= G_0^{-1} - G_0^{-1} A_1^\top \left(I + A_1 G_0^{-1} A_1^\top\right)^{-1} A_1 G_0^{-1}. \end{aligned}$$