Least Squares Solutions and the QR Factorization

Linear Algebra
MATH 2076

UNIVERSITY OF Cincinnati

Least Squares Solution to a System of Linear Equations

A vector $\hat{\mathbf{x}}$ is a *least squares solution* to $A\vec{x} = \vec{b}$ provided for any $\vec{x}$,

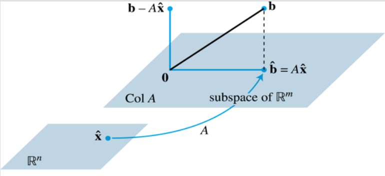
$$\|A\hat{\mathbf{x}} - \vec{b}\| \leq \|A\vec{x} - \vec{b}\| .$$

Here, when A is $m \times n$, $\vec{x}$ is any vector in $\mathbb{R}^n$.

The vector $\hat{\mathbf{b}} = \text{Proj}_{\mathcal{CS}(A)}(\vec{b})$ lies in $\mathcal{CS}(A)$ and is nearest/closest to $\vec{b}$, so any solution $\hat{\mathbf{x}}$ to

$$A\vec{x} = \hat{\mathbf{b}}$$

is a least squares solution.



The least-squares solution $\hat{\mathbf{x}}$ is in $\mathbb{R}^n$.

To find a least squares solution to $A\vec{x} = \vec{b}$:

- Calculate the orthogonal projection $\hat{\mathbf{b}} = \text{Proj}_{\mathcal{CS}(A)}(\vec{b})$. Work!
- Solve $A\vec{x} = \hat{\mathbf{b}}$. Not difficult, right?

Using Geometry to Get a Least Squares Solution

We want to find a solution $\hat{\mathbf{x}}$ to $A\vec{x} = \hat{\mathbf{b}} = \text{Proj}_{\mathcal{CS}(A)}(\vec{b})$.

Recall that $\vec{b} = \hat{\mathbf{b}} + \vec{z}$ where $\vec{z}$ is orthogonal to $\mathcal{CS}(A)$. Thus, $\vec{z}$ lies in the orthogonal complement $\left(\mathcal{CS}(A)\right)^{\perp} = \mathcal{NS}(A^T)$. So,

$$A^T\vec{b} - A^T A\,\hat{\mathbf{x}} = A^T(\vec{b} - A\,\hat{\mathbf{x}}) = A^T(\vec{b} - \hat{\mathbf{b}}) = A^T\vec{z} = \vec{0}.$$

That is,

$$A\,\hat{\mathbf{x}} = \hat{\mathbf{b}} \iff A^T A\,\hat{\mathbf{x}} = A^T\vec{b}.$$

Any solution $\hat{\mathbf{x}}$ to $\boxed{A^T A\vec{x} = A^T\vec{b}}$ is a least squares solution to $A\vec{x} = \vec{b}$.

When will we get a unique least squares solution?

Unique Solutions to $A\vec{x} = \vec{b}$

Recall the Solution Set Trichotomy. What does this tell us?

For which vectors $\vec{b}$ do we know that $A\vec{x} = \vec{b}$ definitely has a solution?

When does $A\vec{x} = \vec{0}$ have a unique solution?
What does this have to do with an REF for A?
What does this mean, if it is true, about the columns of A?

What can we "do" with A if its columns are linearly independent?

If A has linearly independent columns, we can factor A as $A = QR$ where. . . .

The QR Factorization

Let A be a matrix with linearly independent columns, say

$$A = \left[\vec{a}_1 \ \vec{a}_2 \ldots \vec{a}_n\right] \text{ where } \vec{a}_j = \mathrm{Col}_j(A) \text{ is in } \mathbb{R}^m.$$

Then $\mathcal{A} = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ is a basis for the column space $\mathcal{CS}(A)$.

Gram-Schmidt $\mathcal{A}$ to get an orthon basis $\mathcal{U} = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ for $\mathcal{CS}(A)$. Then $\boxed{A = QR}$ where $Q = \left[\vec{u}_1 \ \vec{u}_2 \ldots \vec{u}_n\right]$ and R is an *invertible* upper triangular matrix.

Since $\mathcal{U}$ is *orthonormal*, $Q^TQ = I$, and therefore $\boxed{R = Q^TA}$.

Moreover, $P = QQ^T$ is the standard matrix for the linear transformation $\mathrm{Proj}_{\mathcal{CS}(A)}$. Right?

Unique Least Squares Solutions to $A\vec{x} = \vec{b}$

For any matrix A, the following are equivalent.

- If $A\vec{x} = \vec{b}$ has a solution, it is unique.
- $A\vec{x} = \vec{0}$ if and only if $\vec{x} = \vec{0}$.
- The columns of A are linearly independent.
- $A = QR$ with $Q^TQ = I$ and R invertible upper triangular.
- A^TA is invertible. $(A^TA = (QR)^TQR = R^TQ^TQR = R^TR$:-)
- $A\vec{x} = \vec{b}$ has a unique least squares solution for every $\vec{b}$.
- $A^TA\vec{x} = A^T\vec{b}$ has the unique solution $\hat{\mathbf{x}} = \left(A^TA\right)^{-1}A^T\vec{b}$.
- $A^TA\vec{x} = A^T\vec{b}$ has the unique solution $\hat{\mathbf{x}} = R^{-1}Q^T\vec{b}$.

One more look at $A\vec{x} = \hat{\mathbf{b}} = \text{Proj}_{\mathcal{CS}(A)}(\vec{b})$

Suppose $A = QR$ with $Q^TQ = I$ and R invertible upper triangular.

Since $P = QQ^T$ is the standard matrix for the LT $\text{Proj}_{\mathcal{CS}(A)}$, $\hat{\mathbf{b}} = \text{Proj}_{\mathcal{CS}(A)}(\vec{b}) = P\vec{b} = QQ^T\vec{b}$. So, we can rewrite $A\vec{x} = \hat{\mathbf{b}}$ as

$$QR\vec{x} = A\vec{x} = \hat{\mathbf{b}} = QQ^T\vec{b}$$

and multiplying by Q^T (and remembering that $Q^TQ = I$) we get

$$R\vec{x} = Q^T\vec{b} \quad \text{with the unique solution} \quad \hat{\mathbf{x}} = R^{-1}Q^T\vec{b}.$$

However, R is already upper triangular, so it is almost always better to just solve $R\vec{x} = Q^T\vec{b}$.

Finding a Least Squares Solution via $A = QR$

Let's find a least squares solution to $\begin{bmatrix} 3 & -6 \\ 4 & -8 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ 7 \\ 2 \end{bmatrix}$.

It's easy to find the QR-factorization for the above coefficient matrix:

$$\begin{bmatrix} 3 & -6 \\ 4 & -8 \\ 0 & 1 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 3 & 0 \\ 4 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 5 & -10 \\ 0 & 1 \end{bmatrix}.$$

Now we want to solve $R\vec{x} = Q^T\vec{b}$ which is

$$\begin{bmatrix} 5 & -10 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 3 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} -1 \\ 7 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}.$$

Thus $\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$. Notice that $\begin{bmatrix} 3 & -6 \\ 4 & -8 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$.