Least squares, Gram-Schmidt solved problems

The examples below are of complexity similar to the problems in Homework 7. Please note that even in these small dimensional cases the arithmetic gets tiresome, hiding the simplicity of the solution by projection:

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|_2, \boldsymbol{Q}\boldsymbol{R}=\boldsymbol{A}, \boldsymbol{R}\boldsymbol{x}=\boldsymbol{Q}^T\boldsymbol{b}.$$

The intent of Homework 7 is to:

- 1. Gain familiarity with the Gram-Schmidt algorithm;
- 2. Understand the steps in solving the least squares problem by projection;
- 3. Understand the necessity of working with symbolic representations of vectors and matrices, leaving the arithmetic to software systems (Matlab/Octave).

Theory primer. The least squares problem is stated as

$$\min_{\boldsymbol{x}\in\mathbb{R}^n}\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|_2$$

and seeks to find the scaling coefficients $\boldsymbol{x} \in \mathbb{R}^n$ required to most closely approximate $\boldsymbol{b} \in \mathbb{R}^m$ by linear combination of the columns of $\boldsymbol{A} = [\begin{array}{ccc} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_n \end{array}] \in \mathbb{R}^{m \times n}$.

The solution is found by orthogonal projection of \boldsymbol{b} onto the column space of \boldsymbol{A} . It is efficient to define an orthonormal basis for the column space of \boldsymbol{A} through the column vectors of \boldsymbol{Q}

$$\boldsymbol{Q} = [\boldsymbol{q}_1 \ \boldsymbol{q}_2 \ \ldots \ \boldsymbol{q}_r], \boldsymbol{Q} \boldsymbol{R} = \boldsymbol{A}_r$$

computed using the Gram-Schmidt algorithm (see below). Read QR = A to state: "linear combinations of the columns of Q can be formed to obtain the column vectors of A". Therefore projecting onto the column space of A is the same as projecting onto the column space of Q and the projection matrix is

$$\boldsymbol{P} = \boldsymbol{Q} \boldsymbol{Q}^T$$

where $Q \in \mathbb{R}^{m \times r}$ is a matrix with r mutually orthogonal columns of unit norm, and $r = \operatorname{rank}(A)$. The matrix $Q = [q_1 \ q_2 \ \dots \ q_r]$ is (see below).



The projection of **b** onto $C(\mathbf{A})$ is a vector **v** that (since it's in the column space of **A**) can also be expressed as $\mathbf{v} = \mathbf{A}\mathbf{x}$

$$v = Pb = Ax$$
.

This leads to

$$QQ^T b = QRx \Rightarrow Rx = Q^T b = y$$

Algorithm 1

Given *n* vectors $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ Initialize $\boldsymbol{q}_1 = \boldsymbol{a}_1, \dots, \boldsymbol{q}_n = \boldsymbol{a}_n, \ \boldsymbol{R} = \boldsymbol{I}_n$ for i = 1 to *n* $r_{ii} = (\boldsymbol{q}_i^T \boldsymbol{q}_i)^{1/2}$; if $|r_{ii}| < \varepsilon$ skip to next *i* $\boldsymbol{q}_i = \boldsymbol{q}_i / r_{ii}$ for j = i+1 to *n* $r_{ij} = \boldsymbol{q}_i^T \boldsymbol{a}_j$; $\boldsymbol{q}_j = \boldsymbol{q}_j - r_{ij} \boldsymbol{q}_i$ end end return $\boldsymbol{Q}, \boldsymbol{R}$

Note that if $\boldsymbol{b} \in C(\boldsymbol{A})$ (\boldsymbol{b} is in the column space of \boldsymbol{A}) then $\boldsymbol{b} = \boldsymbol{Q}\boldsymbol{y}$ and

$$Pb = QQ^TQy = Qy = b$$
,

stating that the projection of **b** is **b** itself. In this case the least squares problem has a solution such that b = Ax exactly.

1. Matrix A with linearly independent columns, b not in the column space

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Solution. Note that the column vectors are orthogonal

$$\boldsymbol{a}_1^T \boldsymbol{a}_2 = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = 0.$$

Remember to first look at a problem before blindly carrying out calculations. In this case Q is found by scaling each column vector by its norm

$$Q = [q_1 \ q_2] = \begin{bmatrix} a_1 & a_2 \\ \|a_1\| & \|a_2\| \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix},$$

and the matrix \boldsymbol{R} is a diagonal matrix containing the norms

$$\boldsymbol{R} = \left[\begin{array}{cc} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{array} \right].$$

Verify:

$$\boldsymbol{Q}\boldsymbol{R} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} = \boldsymbol{A}$$

Now that the QR factorization has been found, here are the steps to solve the least squares problem.

a) Find the vector $\boldsymbol{y} = \boldsymbol{Q}^T \boldsymbol{b}$

$$\boldsymbol{y} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

b) Solve the system $\mathbf{R}\mathbf{x} = \mathbf{y}$

$$\boldsymbol{R}\boldsymbol{x} = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \boldsymbol{x} = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \boldsymbol{y}$$

Find

$$\boldsymbol{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \boldsymbol{v} = \boldsymbol{A} \boldsymbol{x} = \begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Computer solution (useful as verification). The above calculations can also be carried out in Matlab/Octave

```
>> A=[1 -1; 0 0; 1 1]; [Q,R]=qr(A,0)

>> Q

\begin{pmatrix} -0.7071 & 0.7071 \\ 0.0 & 0.0 \\ -0.7071 & -0.7071 \end{pmatrix}

>> R

\begin{pmatrix} -1.4142 & -3.3307e - 16 \\ 0.0 & -1.4142 \end{pmatrix}

>>
```

Note that Matlab/Octave returned orthogonal vectors in the opposite orientation, highlighting that there are multiple ways to orthonormalize the columns of A.

In Matlab/Octave the least squares solution is returned by the backslash operator

```
>> b=[1; 1; 1]; x=A\b; disp(x)

1

0

>>
```

2. Matrix A with linearly independent columns, b in the column space

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}.$$

Solution. Note that \boldsymbol{b} is a simple scaling of the first column of \boldsymbol{A}

$$\boldsymbol{b} = 2\boldsymbol{a}_1 = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \boldsymbol{A} \boldsymbol{x} \Rightarrow \boldsymbol{x} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

This example again highlights the need to first look at a problem before blindly carrying out calculations.

3. Matrix \boldsymbol{A} with linearly independent columns, \boldsymbol{b} not in the column space

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}.$$

Solution. In this case the Gram-Schmidt algorithm is applied to find the QR factorization of A. Here are the GS steps:

GS1. Find \boldsymbol{q}_1

$$r_{11} = \|\boldsymbol{a}_1\| = \sqrt{3}, \, \boldsymbol{q}_1 = \frac{\boldsymbol{a}_1}{r_{11}} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1\\1\\1 \end{bmatrix}$$

GS2. Find q_2 . This is carried out in two sub-steps:

a) Subtract from \boldsymbol{a}_2 its component along the previously determined direction \boldsymbol{q}_1

$$\boldsymbol{w} = \boldsymbol{a}_{2} - (\boldsymbol{q}_{1}^{T} \boldsymbol{a}_{2}) \, \boldsymbol{q}_{1} = \boldsymbol{a}_{2} - r_{12} \, \boldsymbol{q}_{1}$$
$$r_{12} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} = \frac{2}{\sqrt{3}}$$
$$\boldsymbol{w} = \boldsymbol{a}_{2} - r_{12} \, \boldsymbol{q}_{1} = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} - \frac{2}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -5 \\ 4 \\ 1 \end{bmatrix}$$

This vector is orthogonal to \boldsymbol{q}_1 . Verify

$$\boldsymbol{q}_{1}^{T}\boldsymbol{w} = \boldsymbol{q}_{1}^{T}(\boldsymbol{a}_{2} - r_{12}\,\boldsymbol{q}_{1}) = r_{12} - r_{12}\,\boldsymbol{q}_{1}^{T}\boldsymbol{q}_{1} = r_{12} - r_{12}\cdot\boldsymbol{1} = 0$$
$$\frac{1}{\sqrt{3}}\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \frac{1}{3} \begin{bmatrix} -5 \\ 4 \\ 1 \end{bmatrix} = \frac{1}{3\sqrt{3}}(-5 + 4 + 1) = 0$$

b) Divide the resulting vector by its norm

$$r_{22} = \|\boldsymbol{w}\| = \frac{\sqrt{42}}{3}, \boldsymbol{q}_2 = \frac{\boldsymbol{w}}{r_{22}} = \frac{1}{\sqrt{42}} \begin{bmatrix} -5 \\ 4 \\ 1 \end{bmatrix}$$

_

_

Verify:

$$\boldsymbol{Q}\boldsymbol{R} = \begin{bmatrix} 1/\sqrt{3} & -5/\sqrt{42} \\ 1/\sqrt{3} & 4/\sqrt{42} \\ 1/\sqrt{3} & 1/\sqrt{42} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 2/\sqrt{3} \\ 0 & \sqrt{42}/3 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix} = \boldsymbol{A}$$

Here are the steps to solve the least squares problem.

a) Find the vector $\boldsymbol{y} = \boldsymbol{Q}^T \boldsymbol{b}$

$$\boldsymbol{y} = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ -5/\sqrt{42} & 4/\sqrt{42} & 1/\sqrt{42} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4/\sqrt{3} \\ 4/\sqrt{42} \end{bmatrix}$$

b) Solve the system Rx = y

$$\boldsymbol{R}\boldsymbol{x} = \begin{bmatrix} \sqrt{3} & 2/\sqrt{3} \\ 0 & \sqrt{42}/3 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 4/\sqrt{3} \\ 4/\sqrt{42} \end{bmatrix} = \boldsymbol{y}$$

Find

$$\boldsymbol{x} = \begin{bmatrix} 24/21 \\ 6/21 \end{bmatrix}, \boldsymbol{v} = \boldsymbol{A}\boldsymbol{x} = \begin{bmatrix} 1 & -1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 24/21 \\ 6/21 \end{bmatrix} = \begin{bmatrix} 6/7 \\ 12/7 \\ 10/7 \end{bmatrix}$$

Computer solution (useful as verification). The above calculations can also be carried out in Matlab/Octave

```
>> A=[1 -1; 1 2; 1 1]; [Q,R]=qr(A,0)

>> Q

\begin{pmatrix} -0.5774 & 0.7715 \\ -0.5774 & -0.6172 \\ -0.5774 & -0.1543 \end{pmatrix}

>> R

\begin{pmatrix} -1.7321 & -1.1547 \\ 0.0 & -2.1602 \end{pmatrix}
```

In Matlab/Octave the least squares solution is returned by the backslash operator

>>