Examples for the *row space in situ* method

These notes supply examples for the *row space in situ* method for solving $Ax = b$, following the preprint by Zimmer[1]. The solutions will be based on this data:

$$A = \begin{bmatrix} 0 & -3i & 0 \\ 2i & 1 & -1 \\ 4i & 2-3i & -2 \end{bmatrix}, \qquad b = \begin{pmatrix} 1 \\ 2i \\ 1+4i \end{pmatrix}$$

By inspection, $m = n = 3$. Also, A has rank 2, by design.

Two equivalent approaches

In the accompanying paper, two related approaches were presented for computing the solution, one without G (a generalized inverse) and one with. For convenience, here are the formulae for the approach without G:

$$\begin{aligned} [A|b] &\to [A'|b'] \\ x_p &= (A')^* b' \\ x_h &= Py \\ P &= 1_n - (A')^* A' \end{aligned}$$

Also, the solution that uses G is formulated as

$$\begin{aligned} [A|1] &\to [A'|M] \\ G &= (A')^* M \\ x_p &= Gb \\ x_h &= Py \\ P &= 1_n - GA \end{aligned}$$

In both cases y is an arbitrary element of C^3.

The orthonormalization steps

This section explicitly shows the steps to orthonormalize A. It follows a *modified* Gram Schmidt approach. In the first approach (i.e., without G), these row operations would be applied to $[A|b]$; in the second approach, they'd be applied to $[A|1]$. In neither case is it necessary to explicitly construct the matrices M_i. They are shown here only for pedagogical purposes. The operations on each row r_i are:

$$\begin{aligned}
&\text{step 1: } r_1 \leftarrow r_1/\|r_1\| &&= r_1/3 \\
&\text{step 2: } r_2 \leftarrow r_2 - \langle r_2, r_1\rangle r_1 &&= r_2 - (i)r_1 \\
&\text{step 3: } r_3 \leftarrow r_3 - \langle r_3, r_1\rangle r_1 &&= r_3 - (3+2i)r_1 \\
&\text{step 4: } r_2 \leftarrow r_2/\|r_2\| &&= r_2/\sqrt{5} \\
&\text{step 5: } r_3 \leftarrow r_3 - \langle r_3, r_2\rangle r_2 &&= r_3 - (2\sqrt{5})r_2
\end{aligned}$$

where $\langle v, w\rangle = w^* v$ and its computation only involves the portion of the row in A. Note that there isn't a step for the normalization of r_3 since in this case its norm is 0. The corresponding matrices for these steps are:

$$M = M_5 M_4 M_3 M_2 M_1$$

$$M_1 = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 1 & 0 & 0 \\ -i & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad M_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -(3+2i) & 0 & 1 \end{bmatrix},$$

$$M_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{5}}{5} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad M_5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2\sqrt{5} & 1 \end{bmatrix}$$

The results: without G

The first approach, which involves orthonormalizing the augmented matrix $[A|b]$, yields

$$[A'|b'] = \begin{bmatrix} 0 & -i & 0 & | & \frac{1}{3} \\ \frac{2}{\sqrt{5}}i & 0 & -\frac{1}{\sqrt{5}} & | & \frac{\sqrt{5}}{3}i \\ 0 & 0 & 0 & | & 0 \end{bmatrix}$$

$$x_p = (A')^* b' = \begin{bmatrix} 0 & -\frac{2}{\sqrt{5}}i & 0 \\ i & 0 & 0 \\ 0 & -\frac{1}{\sqrt{5}} & 0 \end{bmatrix} \begin{pmatrix} \frac{1}{3} \\ \frac{\sqrt{5}}{3}i \\ 0 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 \\ i \\ -i \end{pmatrix}$$

The null space projection operator is

$$P = I_3 - (A')^* A' = \frac{1}{5} \begin{bmatrix} 1 & 0 & -2i \\ 0 & 0 & 0 \\ 2i & 0 & 4 \end{bmatrix}$$

As noted earlier, the homogeneous solution is formed as $x_h = Py$, for arbitrary y. Setting $y = (y_1, y_2, y_3)^T$, where each entry is arbitrary, the result is

$$x_h = \frac{1}{5}(y_1 - 2iy_3)\begin{pmatrix}1\\0\\2i\end{pmatrix} = \alpha\begin{pmatrix}1\\0\\2i\end{pmatrix}$$

where α is an arbitrary element in $\mathcal{C}$. Noteworthy is that the nullity of A is one, which corresponds to the above parametrization of x_h requiring only one vector. The reader should verify that $Ax_p = b$ and $Ax_h = 0$.

The results: with G

The same data for A,b is used here as was used in the previous section. Also, the same orthonormalization steps are implemented, except now they are done on $[A|1]$, as opposed to $[A|b]$. As a reminder, it is *not* necessary to explicitly construct the matrices M_s for the orthonormalization steps; it is only necessary to implement those row operations. The final A' and M matrices are found after step 5 to be:

$$[A'|M] = \begin{bmatrix} 0 & -i & 0 & | & \frac{1}{3} & 0 & 0 \\ \frac{2}{\sqrt{5}}i & 0 & -\frac{1}{\sqrt{5}} & | & -\frac{\sqrt{5}}{15}i & \frac{\sqrt{5}}{5} & 0 \\ 0 & 0 & 0 & | & -1 & -2 & 1 \end{bmatrix}$$

The generalized inverse G is

$$G = (A')^* M = \frac{1}{15}\begin{bmatrix} -2 & -6i & 0 \\ 5i & 0 & 0 \\ i & -3 & 0 \end{bmatrix}$$

The reader is encouraged to verify that computing

$$x_p = Gb$$
$$P = I_3 - GA$$

produces the same results for x_p, P as found in the previous section.

Online case

In the following, the same data is used to illustrate the online case for the first variation. Because the data (for A,b) arrives one row at a time, the version of orthonormalization used is *classical* Gram-Schmidt (CGS). In particular the order of row operations will be:

Row 1 arrives

step 1: $r_1 \leftarrow r_1/3$

Row 2 arrives

step 2: $r_2 \leftarrow r_2 - (i)r_1$

step 3: $r_2 \leftarrow r_2/\sqrt{5}$

Row 3 arrives

step 4: $r_3 \leftarrow r_3 - (3+2i)r_1$

step 5: $r_3 \leftarrow r_3 - (2\sqrt{5})r_2$

Note that in this case row # 3 isn't normalized, as its norm is zero. The pattern to be aware of here is that after the i-th step, the i-th row will no longer change. This means that it can be used toward forming the solution x. Thus, if the data is slow to arrive, steps 1 and 2 can be done while waiting. When the final row arrives (i.e., row 3), the last bit of computation can then be done (i.e., steps 4 and 5). This is why it's called an *online* computation. With this in mind, the solution is rewritten using a column-row expansion

$$x_p = (A')^* b' = \sum_{i=1}^{m} x_p^{(i)}$$

where

$$x_p^{(i)} = \text{Col}_i[(A')^*]\ b'_i$$

and Col_i signifies the i-th column.

In the expressions below, the "input data" denotes the data that has just arrived. (Double-hyphens in a matrix mean that no data has been entered there yet.) Also, the "intermediate results" are how A, b, and x_p appear following the i-th update.

$i = 1$ ———————————————

input data:

$$\text{Row}_1(A) = (0, -3i, 0)$$
$$\text{Row}_1(b) = (1)$$

intermediate results:

$$A' = \begin{bmatrix} 0 & -i & 0 \\ -- & -- & -- \\ -- & -- & -- \end{bmatrix}, \quad b' = \begin{bmatrix} \frac{1}{3} \\ -- \\ -- \end{bmatrix},$$

$$x_p^{(1)} = \mathrm{Col}_1[(A')^*]\ b'_1 = \begin{bmatrix} 0 \\ \frac{1}{3}i \\ 0 \end{bmatrix}$$

$i = 2$ ————————————————

input data:

$$\mathrm{Row}_2(A) = (2i, 1, -1)$$
$$\mathrm{Row}_2(b) = (2i)$$

intermediate results:

$$A' = \begin{bmatrix} 0 & -i & 0 \\ \frac{2\sqrt{5}}{5}i & 0 & -\frac{\sqrt{5}}{5} \\ -- & -- & -- \end{bmatrix}, \quad b' = \begin{bmatrix} \frac{1}{3} \\ \frac{\sqrt{5}}{3}i \\ -- \end{bmatrix},$$

$$x_p^{(2)} = \mathrm{Col}_2[(A')^*]\ b'_2 = \begin{bmatrix} \frac{2}{3} \\ 0 \\ -\frac{1}{3}i \end{bmatrix}$$

$i = 3$ ————————————————

input data:

$$\mathrm{Row}_3(A) = (4i, 2-3i, -2)$$
$$\mathrm{Row}_3(b) = (1+4i)$$

intermediate results:

$$A' = \begin{bmatrix} 0 & -i & 0 \\ \frac{2\sqrt{5}}{5}i & 0 & -\frac{\sqrt{5}}{5} \\ 0 & 0 & 0 \end{bmatrix}, \quad b' = \begin{bmatrix} \frac{1}{3} \\ \frac{\sqrt{5}}{3}i \\ 0 \end{bmatrix},$$

$$x_p^{(3)} = \mathrm{Col}_3[(A')^*]\ b'_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The particular solution follows from adding all the updates, giving

$$x_p = x_p^{(1)} + x_p^{(2)} + x_p^{(3)} = \frac{1}{3}\begin{bmatrix} 2 \\ i \\ -i \end{bmatrix}$$

which is the same as found earlier. As a final point, note that it is trivial to repeat the (CGS) orthonormalization step for each i during this online computation. Doing so would benefit its numerical accuracy.

References

[1] M.F. Zimmer, Two direct solvers for a system of linear equations. *arXiv preprint*, arXiv:1611.06633, 2020.

Michael Zimmer
zim@neomath.com