Problem 1

The solution is provided in a python script.

Note that there are two versions available:

1. ldu1.py
   This implements the LDU factorization in the naive way with the method as described in the lecture.

2. ldu.py
   This implements the LDU factorization of the matrix using the Doolittle's method.[1]

The solutions are implemented on example matrices which are available in the python script. An example output is as shown in Fig.1 for

$$\mathbf{A} = \begin{pmatrix} 10 & 9 & 2 \\ 5 & 3 & 2 \\ 2 & 2 & 2 \end{pmatrix}$$

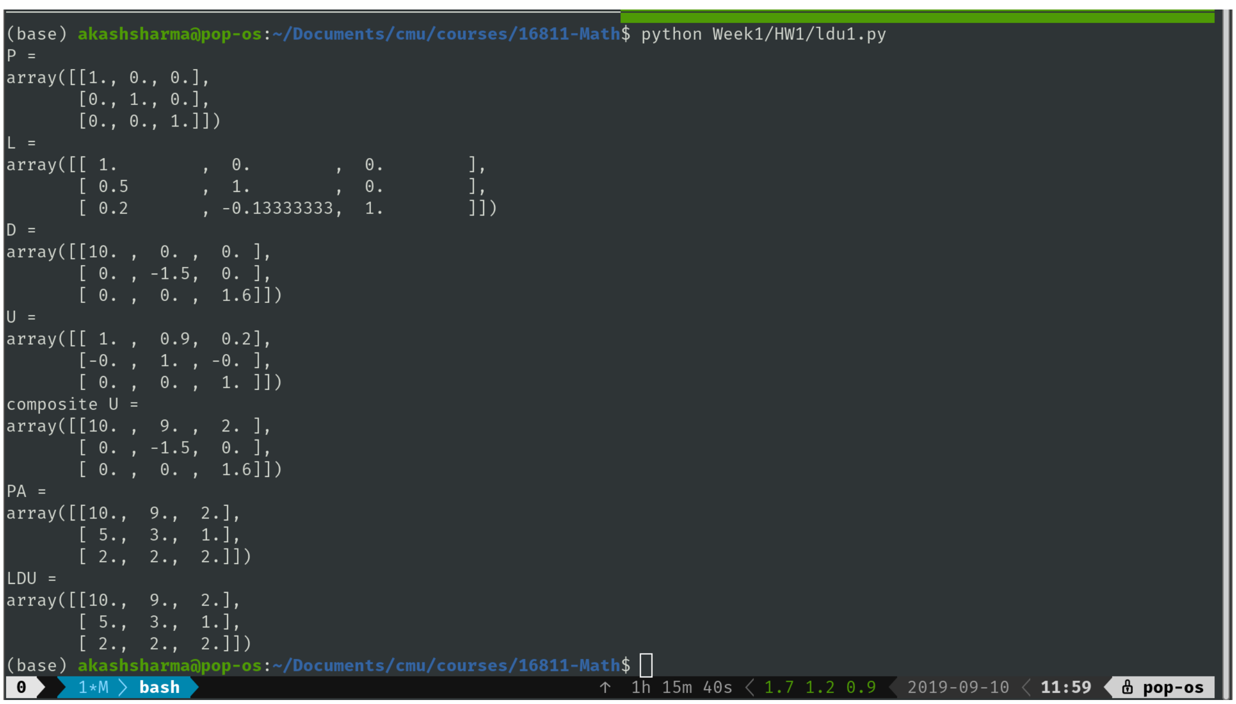

Figure 1: Example solution for Problem 1

Problem 2

I have used the `linalg` functions from the python scipy library, particularly `scipy.linalg.lu` and `scipy.linalg.svd` functions to obtain the LDU and SVD decomposition of the matrices respectively.

1. a. The $PA = LDU$ decomposition for the matrix $\mathbf{A_1}$ is as given below:

$$\overset{P}{\begin{pmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{pmatrix}}\overset{A}{\begin{pmatrix}10 & 9 & 2\\5 & 3 & 1\\2 & 2 & 2\end{pmatrix}} = \overset{L}{\begin{pmatrix}1 & 0 & 0\\0.5 & 1 & 0\\0.2 & -0.13\overline{3} & 1\end{pmatrix}}\overset{D}{\begin{pmatrix}10 & 0 & 0\\0 & -1.5 & 0\\0 & 0 & 1.6\end{pmatrix}}\overset{U}{\begin{pmatrix}1 & 0.9 & 0.2\\0 & 1 & 0\\0 & 0 & 1\end{pmatrix}}$$

   b. The SVD decomposition for the matrix $\mathbf{A_1}$ is as given below:

$$\overset{U}{\begin{pmatrix}-0.8991 & 0.1788 & 0.3997\\-0.3861 & 0.1066 & -0.9162\\-0.2064 & -0.9781 & -0.0268\end{pmatrix}}\overset{\Sigma}{\begin{pmatrix}15.1186 & 0 & 0\\0 & 1.5362 & 0\\0 & 0 & 1.0334\end{pmatrix}}\overset{V^\top}{\begin{pmatrix}-0.7497 & -0.6391 & -0.1718\\0.2376 & -0.0176 & -0.9712\\-0.6177 & 0.7689 & -0.1651\end{pmatrix}}$$

2. a. The $PA = LDU$ decomposition for matrix $\mathbf{A_2}$ is as given below:

$$\overset{P}{\begin{pmatrix}1 & 0 & 0 & 0\\0 & 1 & 0 & 0\\0 & 0 & 1 & 0\\0 & 0 & 0 & 1\end{pmatrix}}\overset{A}{\begin{pmatrix}16 & 16 & 0 & 0\\4 & 0 & -2 & 0\\0 & 1 & -1 & 0\\0 & 0 & 0 & 1\\0 & 0 & 1 & 1\end{pmatrix}} = \overset{L}{\begin{pmatrix}1 & 0 & 0 & 0\\0.25 & 1 & 0 & 0\\0 & -0.25 & 1 & 0\\0 & 0 & 0 & 1\\0 & 0 & 0.667 & 1\end{pmatrix}}\overset{D}{\begin{pmatrix}16 & 0 & 0 & 0\\0 & -4 & 0 & 0\\0 & 0 & -1.5 & 0\\0 & 0 & 0 & 1\end{pmatrix}}\overset{U}{\begin{pmatrix}1 & 1 & 0 & 0\\0 & 1 & 0.5 & 0\\0 & 0 & 1 & 0\\0 & 0 & 0 & 1\end{pmatrix}}$$

   b. The SVD decomposition for matrix $\mathbf{A_2}$ is as given below:

$$\overset{U}{\begin{pmatrix}-0.9914 & 0.1250 & 0 & 0.0268 & -0.0265\\-0.1268 & -0.9728 & 0.1589 & -0.0326 & 0.1060\\-0.0311 & -0.0136 & -0.5399 & -0.7265 & -0.4239\\0 & 0.0173 & 0.3936 & -0.6638 & -0.6358\\0 & 0.1937 & 0.7269 & -0.1729 & 0.6358\end{pmatrix}}\overset{\Sigma}{\begin{pmatrix}22.8186 & 0 & 0 & 0\\0 & 3.4932 & 0 & 0\\0 & 0 & 1.6873 & 0\\0 & 0 & 0 & 1.1227\\0 & 0 & 0 & 0\end{pmatrix}}\overset{V^\top}{\begin{pmatrix}-0.7174 & -0.6965 & 0.0125 & 2.4107\times10^{-5}\\-0.5414 & 0.5687 & 0.6163 & -0.0604\\0.3482 & -0.3486 & 0.5624 & 0.6641\\-0.2664 & -0.2645 & 0.5511 & -0.7452\end{pmatrix}}$$

3. a. The $PA = LDU$ decomposition for matrix $\mathbf{A_3}$ is as given below:

$$\overset{P}{\begin{pmatrix}1 & 0 & 0\\0 & 0 & 1\\0 & 1 & 0\end{pmatrix}}\overset{A}{\begin{pmatrix}10 & 6 & 4\\5 & 3 & 2\\1 & 1 & 0\end{pmatrix}} = \overset{L}{\begin{pmatrix}1 & 0 & 0\\0.1 & 1 & 0\\0.5 & 0 & 1\end{pmatrix}}\overset{D}{\begin{pmatrix}10 & 0 & 0\\0 & 0.4 & 0\\0 & 0 & 0\end{pmatrix}}\overset{U}{\begin{pmatrix}1 & 0.6 & 0.4\\0 & 1 & -1\\0 & 0 & 1\end{pmatrix}}$$

   b. The SVD decomposition for the matrix $\mathbf{A_3}$ is as given below:

$$\overset{U}{\begin{pmatrix}-0.8905 & -0.0840 & -0.4472\\-0.4452 & -0.0419 & 0.8944\\-0.0939 & 0.9956 & 0\end{pmatrix}}\overset{\Sigma}{\begin{pmatrix}13.845 & 0 & 0\\0 & 0.5595 & 0\\0 & 0 & 0\end{pmatrix}}\overset{V^\top}{\begin{pmatrix}-0.8107 & -0.4892 & -0.3216\\-0.0967 & 0.6538 & -0.7505\\0.5774 & -0.5774 & -0.5774\end{pmatrix}}$$

Problem 3

(a) The system provided has an exact solution since the matrix $\mathbf{A}$ is non-singular and square. By inspection it is also clear that the $\mathbf{b}$ also lies in the column space of $\mathbf{A}$. Therefore we solve this using the LDU decomposition, which is given in solution for Problem 2.1(a)

$$\mathbf{Ly} = \mathbf{Pb}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 0.2 & -0.133 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ 2 \\ 4 \end{pmatrix} \implies \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -2 \\ 3 \\ 4.8 \end{pmatrix}$$

$$\mathbf{Ux} = \mathbf{D^{-1}y}$$

$$\begin{pmatrix} 1 & 0.9 & 0.2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{10} & 0 & 0 \\ 0 & \frac{-1}{1.5} & 0 \\ 0 & 0 & \frac{1}{1.6} \end{pmatrix} \begin{pmatrix} -2 \\ 3 \\ 4.8 \end{pmatrix} \implies \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$$

now, by inspection it is evident that b spans the column space of $\mathbf{A}$ and the coefficients are the elements of $\mathbf{x}$. Therefore $\mathbf{x}$ is an exact solution to the system.

$$x_1 \begin{pmatrix} 10 \\ 5 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 9 \\ 3 \\ 2 \end{pmatrix} + x_3 \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \\ 4 \end{pmatrix}$$

(b) In this case, $\mathbf{A}$ is a singular square matrix, since $row_1 = 2 * row_2$. Further the rank of the matrix is 2, as is evident from the k dimensions of the Σ matrix from Problem 2.3(b).

To find whether $\mathbf{b}$ is in the column space of $\mathbf{A}$ we find the row-echelon form of the augmented matrix as below:

$$\left(\begin{array}{ccc|c} 10 & 6 & 4 & 2 \\ 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

We can find c_1, c_2 and c_3 to obtain b from the above matrix, thus $\mathbf{b}$ is in the column space of $\mathbf{A}$. Therefore the SVD solution of the system is the exact solution.

The SVD solution of the system is given below as:

$$\overline{\mathbf{x}} = \mathbf{V} \cdot \frac{\mathbf{1}}{\mathbf{\Sigma}} \cdot \mathbf{U}^{\top}.\mathbf{b}$$

$$\overline{\mathbf{x}} = \begin{pmatrix} -0.8107 & -0.0967 & 0.5774 \\ -0.4892 & 0.6538 & -0.5774 \\ -0.3216 & -0.7505 & -0.5774 \end{pmatrix} \begin{pmatrix} \frac{1}{13.845} & 0 & 0 \\ 0 & \frac{1}{0.5595} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -0.8905 & -0.4452 & -0.0939 \\ -0.0840 & -0.0419 & 0.9956 \\ -0.4472 & 0.8944 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$$

$$\overline{\mathbf{x}} = \begin{pmatrix} 0.3332 \\ -1.3333 \\ 1.6666 \end{pmatrix}$$

Now the null space of the matrix $\mathbf{A}$ is x_n and is given by

$$\mathbf{x_n} = \begin{pmatrix} c_1 \\ -c_1 \\ -c_1 \end{pmatrix}$$

where c_1 could be any constant.

Thus, there are many exact solutions to $\mathbf{Ax} = \mathbf{b}$ and they are of the form $\overline{\mathbf{x}} + \mathbf{x_n}$, and $\overline{\mathbf{x}}$ is the *SVD* solution of the system, which is in the row space of A.

Finally, to verify the solution, if we compute $|\mathbf{A}\overline{\mathbf{x}} - \mathbf{b}|$, we obtain 0, implying that the solution is an exact solution.

(c) As in Problem 3(b) the matrix $\mathbf{A}$ is singular, and has a rank of 2. We reduce the system to the row-echelon form to find whether b is in the column space:

$$\left(\begin{array}{ccc|c} 10 & 6 & 4 & 1 \\ 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

Here, we cannot express $\mathbf{b}$, as a linear combination of $\mathbf{A}$. We can deduce that, $\mathbf{b}$ is not in column space of $\mathbf{A}$. The *SVD* solution is the approximate solution of the system. i.e.,

$$\overline{\mathbf{x}} = \mathbf{V} \cdot \frac{\mathbf{1}}{\mathbf{\Sigma}} \cdot \mathbf{U}^\top.\mathbf{b}$$

Repeating the matrix multiplication as in Problem 3(b) we obtain $\overline{\mathbf{x}}$

$$\overline{\mathbf{x}} = \begin{pmatrix} 0.3332 \\ -1.3333 \\ 1.6666 \end{pmatrix}$$

Here $\overline{\mathbf{x}}$ minimizes the least squares distance $||\mathbf{Ax} = \mathbf{b}||$ and the set of least squares solution for this system would be of form $\overline{\mathbf{x}} + \mathbf{x_n}$ where

$$\mathbf{x_n} = \begin{pmatrix} c_1 \\ -c_1 \\ -c_1 \end{pmatrix}$$

as in Problem 3(b)

We can show that on calculating $|\mathbf{Ax} - \mathbf{b}|$ we get a minimum residue as opposed to 0 in the previous Problem 3(b), which verifies the solution.

Problem 4

(a) For the given matrix $\mathbf{A} = \mathbf{I} - \mathbf{u}\mathbf{u}^\top$, consider transformation for the vector $\mathbf{u}$, now the transformation can be given as

$$\mathbf{A}\mathbf{u} = (\mathbf{I} - \mathbf{u}\mathbf{u}^\top).\mathbf{u} = \mathbf{u} - \mathbf{u}.(\mathbf{u}^\top\mathbf{u}) = \mathbf{u} - \mathbf{u} = 0$$

since $\mathbf{u}^\top\mathbf{u} = 1$, as $\mathbf{u}$ is an unit vector.

Now, consider the vector $\mathbf{v} \perp \mathbf{u}$

$$\mathbf{A}\mathbf{v} = (\mathbf{I} - \mathbf{u}\mathbf{u}^\top).\mathbf{v} = \mathbf{u} - \mathbf{u}.(\mathbf{u}^\top\mathbf{v}) = \mathbf{u} - \mathbf{0} = \mathbf{u}$$

since $\mathbf{u}^\top\mathbf{v} = 0$, as dot product between orthogonal vectors = 0

Now, any vector $\mathbf{x}$ could be represented by an orthogonal basis set = $\{u_1, u_2, \ldots u_n\}$ as

$$\mathbf{x} = \sum_{i=1}^{n}(x.u_i)u_i$$

now if we assume one of the basis vectors $\mathbf{u_1}$ to be along/parallel $\mathbf{u}$, then $\mathbf{A}$ transforms vector X in such away that only the components $\mathbf{u}^\perp = \{u_2, u_3, \ldots, u_n\}$ would remain. Geometrically this would be a projection of vector $\mathbf{x}$ onto the hyper-plane $\mathbf{u}^\perp$.

(b) We know for any matrix A:

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

λ is the eigenvalue of the vector and $\mathbf{v}$ forms the eigenvector corresponding the same. As in previous section, it is evident that Av = 0, when v is parallel to u, implying **eigenvalue = 0** and similarly, when v is perpendicular to u, Av = v, implying repeated **eigenvalues = 1**.

(c) By definition, the null space of a matrix is the space of vectors, which result in $\mathbf{A}\mathbf{x} = 0$, in this case, as in part (a), $\mathbf{A}\mathbf{u} = 0$, this means that all vectors in $span(\mathbf{u})$, lie in the null space of $\mathbf{A}$

(d) Using similar analysis to part (a), we can see that for u:

$$\mathbf{A}^2\mathbf{u} = (\mathbf{I} - \mathbf{u}\mathbf{u}^\top)(\mathbf{I} - \mathbf{u}\mathbf{u}^\top)\mathbf{u} = (\mathbf{I} - \mathbf{u}\mathbf{u}^\top)0 = 0$$

and for $\mathbf{v} \perp \mathbf{u}$

$$\mathbf{A}^2\mathbf{v} = (\mathbf{I} - \mathbf{u}\mathbf{u}^\top)(\mathbf{I} - \mathbf{u}\mathbf{u}^\top)\mathbf{v} = (\mathbf{I} - \mathbf{u}\mathbf{u}^\top)\mathbf{v} = \mathbf{v}$$

It is evident, that there is no change in transformation of vector v or u, on applying the transformation twice, and this implies that

$$\mathbf{A}^2 = \mathbf{A}$$

Additionally, we can also show that

$$\begin{aligned}\mathbf{A^2} &= (\mathbf{I} - \mathbf{u}\mathbf{u}^\top)(\mathbf{I} - \mathbf{u}\mathbf{u}^\top)\\ &= \mathbf{I} - \mathbf{2}(\mathbf{u}\mathbf{u}^\top) + \mathbf{u}(\mathbf{u}^\top\mathbf{u})\mathbf{u}^\top\\ &= \mathbf{I} - \mathbf{2}(\mathbf{u}\mathbf{u}^\top) + \mathbf{u}\mathbf{u}^\top\\ &= \mathbf{I} - \mathbf{u}\mathbf{u}^\top = \mathbf{A}\end{aligned}$$

Problem 5

This question requests an algorithm to find the best/optimal rigid body transformation between two point sets, $\mathcal{P} = \{p_1, p_2, \dots p_n\}$ and $\mathcal{Q} = \{q_1, q_2, \dots q_n\}$ which can be formalized as[2][3]:

$$F = (\mathbf{R}, \mathbf{t}) = \underset{R,t}{\operatorname{argmin}} \sum_{i=1}^{n} ||(Rp_i + t) - q_i||^2 \tag{1}$$

We can find the optimal translation between the two points by taking the derivative w.r.t t, equating the derivative to 0, and finding t.

$$\begin{aligned}\frac{\partial F}{\partial t} &= \sum_{i=1}^{n} 2((Rp_i + t) - q_i)\\ &= 2R\sum_{i=1}^{n} p_i + 2t \cdot n + 2\sum_{i=1}^{n} q_i\\ \implies t &= \frac{1}{n}\sum_{i=1}^{n} q_i - R\frac{1}{n}\sum_{i=1}^{n} p_i\\ t &= \overline{q} - R \cdot \overline{p}\end{aligned}$$

where $\overline{p} = \frac{1}{n}\sum_{i=1}^{n} p_i$ and $\overline{q} = \frac{1}{n}\sum_{i=1}^{n} q_i$

Replacing the value of t in equation 1, we now obtain the minimization as:

$$\begin{aligned}R &= \underset{R}{\operatorname{argmin}} \sum_{i=1}^{n} ||R(p_i - \overline{p}) - (q_i - \overline{q})||^2\\ &= \underset{R}{\operatorname{argmin}} \sum_{i=1}^{n} ||Rx_i - y_i||^2\end{aligned} \tag{2}$$

Now, we aim to find the optimal rotation, we can show from [2] after simplifying the terms that minimizing R can be shown as maximizing the trace of the diagonal matrix $tr(Y^\top RX)$

We compute the SVD decomposition of the covariance matrix $\mathbf{XY}^\top$ to obtain the orthogonal decomposition, i.e.,

$$\mathbf{XY}^\top = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^\top \tag{3}$$

We know that for matrices, $tr(AB) = tr(BA)$, therefore we can write

$$tr(\mathbf{Y}^\top \mathbf{RX}) = tr(\mathbf{RXY}^\top)$$

We need to now maximize $tr(\mathbf{RXY}^\top) = tr(\mathbf{RS}) = tr(\mathbf{RU} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^\top) = tr(\mathbf{\Sigma V}^\top \mathbf{RU})$. Note that $\mathbf{V}^\top, \mathbf{R}, \mathbf{U}$ are orthogonal matrices, which implies that the columns of each of these matrices are required to be orthonormal. Therefore, $\mathbf{M} = \mathbf{V}^\top \mathbf{RU}$ also has orthonormal matrices, further now since we need to find argmax($tr(\mathbf{\Sigma M})$), we note that, we require M to be an identity matrix.

Therefore, we can conclude that

$$\mathbf{M} = \mathbf{I} = \mathbf{V}^\top \mathbf{RU} \implies \mathbf{R} = \mathbf{V} \cdot \mathbf{U}^\top \tag{4}$$

The algorithm is given as below

- Compute the centroid of the point sets through the formula as below

$$\overline{p} = \frac{1}{N}\sum_{i=1}^{N} p_i, \qquad \overline{q} = \frac{1}{N}\sum_{i=1}^{N} q_i$$

- Compute the centered vectors as:

$$x_i = p_i - \overline{p}, \qquad y_i = q_i - \overline{q}$$

- Compute the co-variance matrix between $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$ and $\mathbf{y} = \{y_1, y_2, \ldots, y_n\}$ as

$$\mathbf{H} = \mathbf{xy}^\top \quad \text{where H forms a 3x3 matrix}$$

- Compute the SVD of $\mathbf{H}$ as $\mathbf{U\Sigma V}^\top$, now the required optimal rotation in the least squared sense for the point set is

$$\mathbf{R} = \mathbf{VU}^\top$$

- Finally compute the optimal translation for the rigid body as:

$$\mathbf{t} = -\mathbf{R}\overline{\mathbf{p}} + \overline{\mathbf{q}}$$

Thought

The above problem can also be thought of in a different way:

We know that SVD decomposes the matrices as orthogonal matrices. In the problem given we are trying to find the least squares closest orthogonal matrix $\mathbf{R}$ which approximates $\mathbf{A}$ an over-constrained system of point transformations from $\mathcal{P}$ to $\mathcal{Q}$.

Intuitively to obtain a rigid body transformation we require that there is no shearing of the object in any of the axes/bases, which effectively amounts to $\Sigma = I$, therefore, we could compute as $\mathbf{R} = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^\top = \mathbf{U} \cdot \mathbf{I} \cdot \mathbf{V}^\top$.

Further looking at the previous proof, it seems that this solution gives us the inverse transform from Q to P!

Discussion citation

For this problem, I have discussed with *Tarasha Khurana* andrewID: `tkhurana@andrew.cmu.edu`, and *Rohit Jena* andrewID: `rjena@andrew.cmu.edu`

References

[1] Doolittle method,
`https://www.geeksforgeeks.org/doolittle-algorithm-lu-decomposition/`

[2] Sorkine-Hornung, Olga and Rabinovich, Michael, *Least-squares rigid motion using svd* Computing, 1, 2017

[3] Arun, K. Somani, Thomas S. Huang, and Steven D. Blostein. *Least-squares fitting of two 3-D point sets* IEEE Transactions on pattern analysis and machine intelligence 5 (1987): 698-700.