CSCD37: Assignment #1

1. Consider the overdetermined linear system:

$$Ax = b, \ a \in \mathbb{R}^{m \times n}, \ b \in \mathbb{R}^m, \ x \in \mathbb{R}^n, \ m > n.$$

In lecture we proved that the best least-squares solution x^* which minimizes $\|Ax - b\|_2$ satisfies the normal equations $A^T A x = A^T b$. Alternatively, we could proceed to find $x*$ as follows: Let

$$S(x) = \|Ax - b\|_2^2$$

be a function of the n variables $x_1, x_2, \ldots, x_n$. A well-known result in calculus states that a necessary condition for S to have a minimum at x^* is that x^* satisfies the system of equations

$$\frac{\partial S}{\partial x_i} = 0, \ i = 1, 2, \ldots, n.$$

Show that this system is exactly the normal equations.

What does the partial derivatives system look like? First, need to break down $Ax - b$.

$$Ax - b = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^n a_{1,i} x_i \\ \sum_{i=1}^n a_{2,i} x_i \\ \vdots \\ \sum_{i=1}^n a_{n,i} x_i \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n a_{1,i} x_i - b_1 \\ \sum_{i=1}^n a_{2,i} x_i - b_2 \\ \vdots \\ \sum_{i=1}^n a_{m,i} x_i - b_m \end{bmatrix}$$

Hence $S(x)$ is

$$\left(\sum_{i=1}^n a_{1,i} x_i - b_1 \right)^2 + \left(\sum_{i=1}^n a_{2,i} x_i - b_2 \right)^2 + \ldots + \left(\sum_{i=1}^n a_{m,i} x_i - b_m \right)^2$$

So for some arbitrary x_k, because the other x_j are held constant, $\frac{\partial S}{\partial x_k} = 0$ will be:

$$2 \left(\sum_{i=1}^n a_{1,i} x_i - b_1 \right) a_{1,k} + 2 \left(\sum_{i=1}^n a_{2,i} x_i - b_2 \right) a_{2,k} + \ldots + 2 \left(\sum_{i=1}^n a_{m,i} x_i - b_m \right) a_{m,k} = 0$$

$$= \left(\sum_{i=1}^n (a_{1,k} a_{1,i} x_i) - a_{1,k} b_1 \right) + \left(\sum_{i=1}^n (a_{2,k} a_{2,i} x_i) - a_{2,k} b_2 \right) + \ldots + \left(\sum_{i=1}^n (a_{m,k} a_{m,i} x_i) - a_{m,k} b_m \right) = 0$$

$$= \sum_{j=1}^m \left[\sum_{i=1}^n (a_{j,k} a_{j,i} x_i) - a_{j,k} b_j \right] = 0$$

Comparing this to the normal equations, what does that look like?

$$A^T A x = \begin{bmatrix} a_{1,1} & a_{2,1} & \cdots & a_{m,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,n} & a_{2,n} & \cdots & a_{m,n} \end{bmatrix} \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^{m} a_{i,1} a_{i,1} & \sum_{i=1}^{m} a_{i,1} a_{i,2} & \cdots & \sum_{i=1}^{m} a_{i,1} a_{i,n} \\ \sum_{i=1}^{m} a_{i,2} a_{i,1} & \sum_{i=1}^{m} a_{i,2} a_{i,2} & \cdots & \sum_{i=1}^{m} a_{i,2} a_{i,n} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{m} a_{i,n} a_{i,1} & \sum_{i=1}^{m} a_{i,n} a_{i,2} & \cdots & \sum_{i=1}^{m} a_{i,n} a_{i,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \begin{bmatrix} x_1 \sum_{i=1}^{m} a_{i,1} a_{i,1} + x_2 \sum_{i=1}^{m} a_{i,1} a_{i,2} + \cdots + x_n \sum_{i=1}^{m} a_{i,1} a_{i,n} \\ x_1 \sum_{i=1}^{m} a_{i,2} a_{i,1} + x_2 \sum_{i=1}^{m} a_{i,2} a_{i,2} + \cdots + x_n \sum_{i=1}^{m} a_{i,2} a_{i,n} \\ \vdots \\ x_1 \sum_{i=1}^{m} a_{i,n} a_{i,1} + x_2 \sum_{i=1}^{m} a_{i,n} a_{i,2} + \cdots + x_n \sum_{i=1}^{m} a_{i,n} a_{i,n} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^{m} \left[(x_1 a_{i,1} a_{i,1}) + (x_2 a_{i,1} a_{i,2}) + \cdots + (x_n a_{i,1} a_{i,n}) \right] \\ \sum_{i=1}^{m} \left[(x_1 a_{i,2} a_{i,1}) + (x_2 a_{i,2} a_{i,2}) + \cdots + (x_n a_{i,2} a_{i,n}) \right] \\ \vdots \\ \sum_{i=1}^{m} \left[(x_1 a_{i,n} a_{i,1}) + (x_2 a_{i,n} a_{i,2}) + \cdots + (x_n a_{i,n} a_{i,n}) \right] \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^{m} \left[\sum_{j=1}^{n} (a_{i,1} a_{i,j} x_j) \right] \\ \sum_{i=1}^{m} \left[\sum_{j=1}^{n} (a_{i,2} a_{i,j} x_j) \right] \\ \vdots \\ \sum_{i=1}^{m} \left[\sum_{j=1}^{n} (a_{i,n} a_{i,j} x_j) \right] \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{m} \left[\sum_{i=1}^{n} (a_{j,1} a_{j,i} x_i) \right] \\ \sum_{j=1}^{m} \left[\sum_{i=1}^{n} (a_{j,2} a_{j,i} x_i) \right] \\ \vdots \\ \sum_{j=1}^{m} \left[\sum_{i=1}^{n} (a_{j,n} a_{j,i} x_i) \right] \end{bmatrix}$$

$$A^T b = \begin{bmatrix} a_{1,1} & a_{2,1} & \cdots & a_{m,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,n} & a_{2,n} & \cdots & a_{m,n} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{m} a_{j,1} b_j \\ \sum_{j=1}^{m} a_{j,2} b_j \\ \vdots \\ \sum_{j=1}^{m} a_{j,m} b_j \end{bmatrix}$$

Putting all this into one equation, $A^T A x = A^T b \iff A^T A x - A^T b = 0$ So

$$A^T A x - A^T b = 0$$

$$\begin{bmatrix} \sum_{j=1}^{m} \left[\sum_{i=1}^{n} (a_{j,1} a_{j,i} x_i) \right] \\ \sum_{j=1}^{m} \left[\sum_{i=1}^{n} (a_{j,2} a_{j,i} x_i) \right] \\ \vdots \\ \sum_{j=1}^{m} \left[\sum_{i=1}^{n} (a_{j,n} a_{j,i} x_i) \right] \end{bmatrix} - \begin{bmatrix} \sum_{j=1}^{m} a_{j,1} b_j \\ \sum_{j=1}^{m} a_{j,2} b_j \\ \vdots \\ \sum_{j=1}^{m} a_{j,m} b_j \end{bmatrix} = 0$$

$$\begin{bmatrix} \sum_{j=1}^{m} \left[\sum_{i=1}^{n} (a_{j,1} a_{j,i} x_i) - a_{j,1} b_j \right] \\ \sum_{j=1}^{m} \left[\sum_{i=1}^{n} (a_{j,2} a_{j,i} x_i) - a_{j,2} b_j \right] \\ \vdots \\ \sum_{j=1}^{m} \left[\sum_{i=1}^{n} (a_{j,n} a_{j,i} x_i) - a_{j,m} b_j \right] \end{bmatrix} = 0$$

Which is equivalent to the arbitrary x_k system shown above for $S(x)$ when expanded to $k \in [1, n]$

2. In CSCC37 we saw how to compute the $PA = LU$ factorization of $A \in \mathbb{R}^{n \times n}$. The algorithm involved pre-multiplying A with $n-1$ permutation matrices $\mathcal{P}_i$ interleaved with $n-1$ Gauss transforms $\mathcal{L}_i$, where $\mathcal{P}_i, \mathcal{L}_i \in \mathbb{R}^{n \times n}$, $i = 1, \dots, n-1$. This transformation triangularizes A, producing the upper-triangular matrix $U \in \mathbb{R}^{n \times n}$. The unit-lower triangular matrix $L \in \mathbb{R}^{n \times n}$ is later computed by a sequence of algebraic manipulations on $\mathcal{P}_i, \mathcal{L}_i$, $i = 1, \dots, n-1$.

   (a) Show how $A \in \mathbb{R}^{m \times n}$, $m > n$ can be triangularized in a similar fashion using permutation matrices and Gauss transforms. Clearly specify the dimension and purpose of each $\mathcal{P}_i$ and $\mathcal{L}_i$ required, and the number of $\mathcal{P}_i$ and $\mathcal{L}_i$ required.

   (**Note:** You *do not* need to show how to compute the final L.)

   In this case of an overdetermined system, the LU factorization has to be defined, as L and U are not both square. L will be a square matrix as to be able to chain multiple multiplications together, but U will be the reduced form of A with its upper n rows forming an upper triangular matrix and its lower rows being 0.

   Given $A \in \mathbb{R}^{m \times n}, m > n$ we can use permutation matrices $\mathcal{P}_i \in \mathbb{R}^{m \times m}$ to place the largest (magnitude) element of each column to the current row that we are eliminating from, similar to the method for square matrices. This is done by swapping the row we wish to eliminate from, to the row with the largest element in that column for each column.

   Not only do the $\mathcal{P}_i$ have to be $m \times m$, but so do $\mathcal{L}_i$ to be able to multiply against the other $\mathcal{P}_i$ and $\mathcal{L}_i$. Each of these, again analogous to the square case, will be used to eliminated everything underneath the kth row for the kth column in the exact same way using the diagonal element times the negative inverse of each other element.

   Since there are only n columns that need to be eliminated, we will need n of both $\mathcal{P}$ and $\mathcal{L}$. Note that this is different from the square case, we need one more of each. This is the case because we are not finished after doing the second last column, there will still be elements below the diagonal since we have a rectangular matrix so one more of each is required to clear the column.

   (b) Given that the $PA = LU$ factorization of $A \in \mathbb{R}^{m \times n}$, $m > n$ exists, **either** show how to use the factorization to solve the least squares problem $\min_x \|Ax - b\|_2$, **or** argue why the $PA = LU$ factorization **cannot** be used.

$PA = LU$ will not work for minimizing the norm. If the reason for QR factorization working is examined, we see that $\|QRx - b\|_2 = \|Q^T[QRx - b]\|_2$ which leads to $\|Ax - b\|_2 = Rx - \tilde{b}$. However, in the case of LU factorization, we cannot multiply the equation by an inverse of L and expect the norm to stay the same, this only worked since Q orthogonal. This means that reducing the equation to $Ux = \tilde{b}$ will change the norm, and thus have no bearing on the original equation, so solving $LUx = \tilde{b}$ does not minimize the 2 norm of the original equation.

(c) Given that the $PA = LU$ factorization of $A \in \mathbb{R}^{m \times n}$, $m > n$ exists, **either** show how to use the factorization to solve the null space problem $\min_x \|Ax - b\|_2$, **or** argue why the $PA = LU$ factorization **cannot** be used. (Recall that in the null space problem, $A \in \mathbb{R}^{m \times n}, m > n$ arises by transposing the original coefficient matrix in the undetermined system.)

The factorization cannot be used since again, the orthogonality of Q is essential in the solution. If we examine the proof for the validity of the QR factorization, and re create it for $PA = LU$ we get,

$$PA = LU$$
$$(PA)^T = U^T L^T$$
$$(PA)^T L^{-1^T} = U^T$$

and here we see that L^{-1^T} is what is being multiplied by

3. Consider the overdetermined linear system:

$$Ax = b, \ a \in \mathbb{R}^{m \times n}, \ b \in \mathbb{R}^m, \ x \in \mathbb{R}^n, \ m > n.$$

In general there is no exact solution to this system and the best we can do to solve the problem is minimize the norm of the residual $Ax - b$. In certain cases, though, there *is* an exact solution and the norm of the residual is zero. State the conditions required for this to happen and give an example.

4. In lecture we proved that the Householder transformation

$$Q = I - 2\frac{vv^T}{v^T v}$$

is both orthogonal and symmetric for any $v \neq 0$, and we *claimed* that if we choose

$$v = x - \|x\|_2 e_1 \tag{1}$$

we have

$$Qx = \left[I - 2\frac{vv^T}{v^T v}\right] x = \|x\|_2 e_1 \tag{2}$$

where e_1 is the first column of the identity matrix. We then used (2) to compute the QR-factorization by embedding a sequence of similar transformations in the identity matrix. Prove that if we choose v as in (1), (2) holds.

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \ v = \begin{bmatrix} x_1 - \|x\|_2 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$v^T v = \begin{bmatrix} x_1 - \|x\|_2 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} x_1 - \|x\|_2, & x_2, & \cdots, & x_n \end{bmatrix}$$

$$= (x_1 - \|x\|_2)^2 + x_2^2 + \cdots + x_n^2$$

$$= \|x\|_2^2 - 2x_1\|x\|_2 + \sum_{i=1}^{n} x_i^2$$

$$= \sum_{i=1}^{n} x_i^2 - 2x_1\|x\|_2 + \sum_{i=1}^{n} x_i^2$$

$$= 2\sum_{i=1}^{n} x_i^2 - 2x_1\|x\|_2$$

$$vv^T = \begin{bmatrix} (x_1 - \|x\|_2)^2 & (x_1 - \|x\|_2)x_2 & \cdots & (x_1 - \|x\|_2)x_n \\ (x_1 - \|x\|_2)x_2 & x_2^2 & \cdots & x_2 x_n \\ \vdots & \vdots & \ddots & \vdots \\ (x_1 - \|x\|_2)x_n & x_2 x_n & \cdots & x_n^2 \end{bmatrix}$$

$$Qx = \left[I - 2\frac{vv^T}{v^T v} \right] x$$

$$= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - \frac{2}{2\sum_{i=1}^{n} x_i^2 - 2x_1\|x\|_2} \begin{bmatrix} (x_1 - \|x\|_2)^2 x_1 + (x_1 - \|x\|_2)x_2^2 + \cdots + (x_1 - \|x\|_2)x_n^2 \\ (x_1 - \|x\|_2)x_2 x_1 + x_2^3 + \cdots + x_2 x_n^2 \\ \vdots \\ (x_1 - \|x\|_2)x_n x_1 + x_2^2 x_n + \cdots + x_n^3 \end{bmatrix}$$

$$= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - \frac{1}{\sum_{i=1}^{n} x_i^2 - x_1\|x\|_2} \begin{bmatrix} (x_1 - \|x\|_2)\left[(x_1 - \|x\|_2)x_1 + \sum_{i=2}^{n} x_i^2\right] \\ x_2\left[(x_1 - \|x\|_2)x_1 + \sum_{i=2}^{n} x_i^2\right] \\ \vdots \\ x_n\left[(x_1 - \|x\|_2)x_1 + \sum_{i=2}^{n} x_i^2\right] \end{bmatrix}$$

$$= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - \frac{1}{\sum_{i=1}^{n} x_i^2 - x_1\|x\|_2} \begin{bmatrix} (x_1 - \|x\|_2)\left[-x_1\|x\|_2 + \sum_{i=1}^{n} x_i^2\right] \\ x_2\left[-x_1\|x\|_2 + \sum_{i=1}^{n} x_i^2\right] \\ \vdots \\ x_n\left[-x_1\|x\|_2 + \sum_{i=1}^{n} x_i^2\right] \end{bmatrix}$$

$$= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - \begin{bmatrix} x_1 - \|x\|_2 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \|x\|_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \text{ As wanted}$$

5. Consider the overdetermined linear system $Ax = b$ where

$$A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \\ 0 & 4 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

(a) Compute the QR-factorization of A using Householder reflections. Show *all* intermediate calculations. For example, the first reflection is given by

$$\mathcal{Q}_1 = I - 2\frac{vv^T}{v^T v}, \quad v = a_1 - \|a_1\|_2 e_1$$

where a_1 is the first column of A. Verify that $Q = \mathcal{Q}_1^T \mathcal{Q}_2^T$ is orthogonal.

$$v = \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix} - \sqrt{3^2 + 4^2}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix} - \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 4 \\ 0 \end{bmatrix}$$

$$v^T v = (-2)^2 + 4^2 = 20$$

$$vv^T = \begin{bmatrix} -2 \\ 4 \\ 0 \end{bmatrix}\begin{bmatrix} -2 & 4 & 0 \end{bmatrix} = \begin{bmatrix} 4 & -8 & 0 \\ -8 & 16 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathcal{Q}_1 = I - 2\frac{vv^T}{v^T v} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{1}{10}\begin{bmatrix} 4 & -8 & 0 \\ -8 & 16 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 3/5 & 4/5 & 0 \\ 4/5 & -3/5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathcal{Q}_1 A = \begin{bmatrix} 3/5 & 4/5 & 0 \\ 4/5 & -3/5 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 4 & 5 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 0 & -3 \\ 0 & 4 \end{bmatrix}$$

$$a_2 = \begin{bmatrix} -3 \\ 4 \end{bmatrix}, \quad v' = a_2 - \|a_2\|_2 e_1 = \begin{bmatrix} -3 \\ 4 \end{bmatrix} - \sqrt{(-3)^2 + (4)^2}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -8 \\ 4 \end{bmatrix}$$

$$v'^T v' = (-8)^2 + 4^2 = 80, \quad v'v'^T = \begin{bmatrix} -8 \\ 4 \end{bmatrix}\begin{bmatrix} -8 & 4 \end{bmatrix} = \begin{bmatrix} 64 & -32 \\ -32 & 16 \end{bmatrix}$$

$$\mathcal{Q}_2 = I - 2\frac{v'v'^T}{v'^T v'} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{40}\begin{bmatrix} 64 & -32 \\ -32 & 16 \end{bmatrix} = \begin{bmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{bmatrix}$$

Expanded to $\mathbb{R}^{3\times 3}$,

$$\mathcal{Q}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3/5 & 4/5 \\ 0 & 4/5 & 3/5 \end{bmatrix}$$

This means that

$$Q = \mathcal{Q}_1^T \mathcal{Q}_2^T = \begin{bmatrix} 3/5 & 4/5 & 0 \\ 4/5 & -3/5 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & -3/5 & 4/5 \\ 0 & 4/5 & 3/5 \end{bmatrix} = \begin{bmatrix} 3/5 & -12/25 & 16/25 \\ 4/5 & 9/25 & -12/25 \\ 0 & 4/5 & 3/5 \end{bmatrix}$$

Also, $R = Q_2 Q_1 A$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -3/5 & 4/5 \\ 0 & 4/5 & 3/5 \end{bmatrix} Q_2 Q_1 A = \begin{bmatrix} 5 & 4 \\ 0 & -3 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 0 & 5 \\ 0 & 0 \end{bmatrix}$$

Verifying that Q is orthogonal,

$$QQ^T = \begin{bmatrix} 3/5 & 4/5 & 0 \\ 4/5 & -3/5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3/5 & 4/5 \\ 0 & 4/5 & 3/5 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3/5 & 4/5 \\ 0 & 4/5 & 3/5 \end{bmatrix} \begin{bmatrix} 3/5 & 4/5 & 0 \\ 4/5 & -3/5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 3/5 & 4/5 & 0 \\ 4/5 & -3/5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3/5 & 4/5 & 0 \\ 4/5 & -3/5 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3/5 & 4/5 & 0 \\ 4/5 & -3/5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3/5 & 4/5 & 0 \\ 4/5 & -3/5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ As wanted}$$

(b) Use the QR-factorization in (a) to solve the least squares problem $\min_x \|Ax - b\|_2$.

$$Ax = b$$
$$QRx = b$$
$$Rx = Q^T b$$

$$Q^T b = \begin{bmatrix} 3/5 & 4/5 & 0 \\ -12/25 & 9/25 & 4/5 \\ 16/25 & -12/25 & 3/5 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 11/5 \\ 66/25 \\ 37/25 \end{bmatrix} \implies \tilde{b}_1 = \begin{bmatrix} 11/5 \\ 66/25 \end{bmatrix}$$

$$R = \begin{bmatrix} 5 & 4 \\ 0 & 5 \\ 0 & 0 \end{bmatrix} \implies \mathcal{R}_1 = \begin{bmatrix} 5 & 4 \\ 0 & 5 \end{bmatrix}$$

$$66/25 = 5x_2^* \implies x_2^* = 66/125$$
$$11/5 = 5x_1^* + 4x_2^* \implies 275/125 = 5x_1^* + 264/125$$
$$\implies 11/125 = 5x_1^*$$
$$\implies x_1^* = 11/625$$
$$\implies x^* = \begin{bmatrix} 11/625 \\ 66/125 \end{bmatrix}$$