## Problem 1 (Problem 3.1)

We proceed by induction.
Fix column $i$, when $j = 1$, we have

$$r_{1i}^{CGS} = q_1^T a_i = q_1^T q_i = r_{1i}^{MGS}$$

Since before going into second for loop, $q_i = a_i$. Now suppose $r_{ji}^{CGS} = r_{ji}^{MGS}$ for all $j < k$. When $j = k$, we have

$$r_{ki}^{CGS} = q_k^T a_i$$

$$r_{ki}^{MGS} = q_k^T q_i = q_k^T \left( a_i - \sum_{j=1}^{k-1} r_{ji} q_j \right) = q_k^T a_i$$

Hence, $r_{ki}^{CGS} = r_{ki}^{MGS}$.
By induction, $r_{ji}^{CGS} = r_{ji}^{MGS}$ for all $j$.
Since the choice of $i$ is arbitrary, it follows that CGS and MGS gives the same result under exact precision.

**Problem 2 (Question 3.3)**

1.
Multiplying out the matrix, we have

$$\begin{cases} r + Ax = b \\ A^T r = 0 \end{cases}$$

Let $x$ be the solution of the above system, then

$$A^T(r + Ax) = A^T b \implies A^T Ax = A^T b$$

Since $x$ satisfy the normal equation, it follows that $x$ is a minimizer of $||Ax - b||_2$

2.
Since $A$ has full column rank, it has a SVD $A = U\Sigma V^T$. With this, we see that

$$\begin{bmatrix} I & A \\ A^T & 0 \end{bmatrix} = \begin{bmatrix} I & U\Sigma V^T \\ V\Sigma^T U^T & 0 \end{bmatrix} = \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} I_m & \Sigma \\ \Sigma^T & 0 \end{bmatrix} \begin{bmatrix} U^T & 0 \\ 0 & V^T \end{bmatrix}$$

We compute the eigenvalue by finding the roots of the characteristic polynomial

$$\begin{aligned} \det \begin{bmatrix} (1-\lambda)I_m & \Sigma \\ \Sigma^T & -\lambda I_n \end{bmatrix} &= \det((1-\lambda)I_m)\det(-\lambda I_n - \Sigma^T(1-\lambda)^{-1}I_m\Sigma) \\ &= \det((1-\lambda)I_m)\det(-\lambda I_n - (1-\lambda)^{-1}\Sigma^T\Sigma) \\ &= (1-\lambda)^m \prod_{i=1}^{n}(\lambda - \frac{\sigma_i^2}{1-\lambda}) = 0 \end{aligned}$$

Where $\sigma_i$ denote the singular values of $A$. Hence, $\lambda = 1$ or

$$\lambda^2 - \lambda - \sigma_i^2 = 0 \implies \lambda = \frac{1}{2}(1 \pm \sqrt{1 + 4\sigma_i^2})$$

Therefore,

$$\lambda_{MAX} = \max_i(\frac{1}{2}(1 \pm \sqrt{1 + 4\sigma_i^2}), 1)$$

$$\lambda_{MIN} = \min_i(\frac{1}{2}(1 \pm \sqrt{1 + 4\sigma_i^2}), 1)$$

And the condition number is $|\lambda_{MAX}/\lambda_{MIN}|$.

3.
We apply Gaussian elimination to transform the original matrix into an identity. We first

clear the second row

$$\begin{bmatrix} I_m & 0 \\ -A^T & I_n \end{bmatrix} \begin{bmatrix} I_m & A \\ A^T & 0 \end{bmatrix} = \begin{bmatrix} I_m & A \\ 0 & -AA^T \end{bmatrix}$$

We then clear the 1,2 block

$$\begin{bmatrix} I_m & A(AA^T)^{-1} \\ 0 & I_n \end{bmatrix} \begin{bmatrix} I_m & A \\ 0 & -AA^T \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ 0 & -AA^T \end{bmatrix}$$

Finally, we adjust the 2,2 block

$$\begin{bmatrix} I_m & 0 \\ 0 & -(AA^T)^{-1} \end{bmatrix} \begin{bmatrix} I_m & 0 \\ 0 & -AA^T \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ 0 & I_n \end{bmatrix}$$

Hence, the inverse is given by

$$\begin{bmatrix} I_m & 0 \\ 0 & -(AA^T)^{-1} \end{bmatrix} \begin{bmatrix} I_m & A(AA^T)^{-1} \\ 0 & I_n \end{bmatrix} \begin{bmatrix} I_m & 0 \\ -A^T & I_n \end{bmatrix} = \begin{bmatrix} I_m - A(A^TA)^{-1}A^T & A(A^TA)^{-1} \\ (A^TA)^{-1}A^T & -(A^TA)^{-1} \end{bmatrix}$$

**Problem 3 (Question 3.4)**

To derive the normal equations, we seek to find $x$ such that the gradient vanishes

$$\begin{aligned}
&\lim_{e\to 0} \frac{(A(x+e)-b)^T C(A(x+e)-b) - (Ax-b)^T C(Ax-b)}{e^T C e} \\
&= \lim_{e\to 0} \frac{2e(A^T CAx - A^T Cb) + e^T AXAe}{e^T Ce} \\
&= \lim_{e\to 0} \frac{2e(A^T CAx - A^T Cb)}{e^T Ce}
\end{aligned}$$

In order for the gradient to vanish, we need

$$A^T CAx - A^T Cb = 0 \implies A^T CAx = A^T Cb$$

This gives the normal equation for weighted least squares.

To derive the formulation corresponding to the previous problem, we write

$$\begin{bmatrix} I & C^{1/2}A \\ A^T C^{1/2} & 0 \end{bmatrix} \begin{bmatrix} r \\ x \end{bmatrix} = \begin{bmatrix} C^{1/2}b \\ 0 \end{bmatrix}$$

**Problem 4 (Problem 3.7)**

Suppose $u, v$ are non-zero.
Since $uv^T$ is a rank one matrix, we can decompose it easily with Given's rotation.
In particular, we can apply a sequence of Givens rotation that cancels the $i$th row using the $n$th row (this is possible each each row is essentially a scalar multiple of another row). This process requires $n-1$ Givens rotations

$$\tilde{A} = G_{n-1}G_{n-2}...G_2G_1(R + uv^T)$$

Since $R$ is upper triangular, $G_iR$ is again an upper triangular matrix.
After apply a sequence of Givens rotation, $uv^T$ becomes an upper triangular matrix except the last row. To cancel the last row, we use the first row to cancel out the last row (using another Given's rotations). Hence, the product

$$Q^{-1} = G_nG_{n-1}G_{n-2}...G_2G_1$$

Is the orthogonal matrix of interest

**Problem 5 (Problem 3.15)**

1. Normal equations

For every $x \in \mathbb{R}^n$, write $x$ as $x = x_1 + \tilde{x}$, where $x_1 \in \text{Im}(A^T), x_2 \in \text{Null}(A)$. Thus

$$||Ax - b||_2 = ||Ax_1 - b||_2 = ||AA^Ty - b||_2$$

For some $y$. Since $AA^T$ has full rank,

$$AA^Ty = b$$

Has a solution. However, note that every vector of the form $A^Ty + \tilde{x}$ minimizes the residual square. Since

$$\dim(\text{Null}(A)) = m - n$$

It follows that the solution space has dimension $m-n$. The one that minimizes the norm is $x = A^T(AA^T)^{-1}b$.

2. QR decomposition

Note that $A^T = QR$ for some orthogonal $Q$ and upper triangular $R$, then the least squares problem becomes

$$R^TQ^Tx = b \implies x = Q(R^T)^{-1}b + \tilde{x}$$

Where $\tilde{x} \in \text{Null}(A)$. The norm of the solution is given by

$$||x||_2^2 = x = ||Q(R^T)^{-1}b + \tilde{x}||_2^2 \geq ||Q(R^T)^{-1}b||_2^2$$

Hence, the solution with minimum norm is $Q(R^T)^{-1}b$

3. SVD

Assume $A = V\Sigma U^T$, then the least squares problem becomes

$$U^Tx = \Sigma^{-1}V^Tb \implies x = U\Sigma^{-1}V^Tb + \tilde{x}$$

Where $\tilde{x} \in \text{Null}(A)$. The norm of the solution is given by

$$||x||_2^2 = ||U\Sigma^{-1}V^Tb + \tilde{x}||_2^2 \geq ||U\Sigma^{-1}V^Tb||_2^2$$

Hence, the solution with minimum norm is $U\Sigma^{-1}V^Tb$

## Problem 6 (Problem 3.17)

We proceed by induction. Note that the base case is true:

$$\begin{aligned} P_1P_2 &= (I - u_2u_2^T)(I - u_1u_1^T) \\ &= I - \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 4u_2^Tu_1 & 2 \end{bmatrix} \begin{bmatrix} u_1^T \\ u_2^T \end{bmatrix} \end{aligned}$$

Suppose the claim is true for $n = k - 1$, then for $n = k$, we have

$$\begin{aligned} P_kP_{k-1}...P_1 &= (I - u_ku_k^T)(I - U_{k-1}T_{k-1}U_{k-1}^T) \\ &= I - \begin{bmatrix} U_{k-1} & u_k \end{bmatrix} \begin{bmatrix} T_{k-1} & 0 \\ 4u_k^TU_{k-1}T_{k-1} & 2 \end{bmatrix} \begin{bmatrix} U_{k-1}^T \\ u_k^T \end{bmatrix} \\ &= I - U_kT_kU_k^T \end{aligned}$$

Hence, the claim is true by induction. The inductive proof provides us an algorithm for finding $T$.

```
T_1 = 1
for i = 2 to k:
    b_i = 2u_i^TU_i-1T_i-1
    T_i = [[T_i-1 0], [b_i, 2]]
```