COT 5615 Math for Intelligent Systems, Fall 2021

Homework 8

A16.2 (25 pts)

The code and running time are shown in Fig.1.

```
for i in 1:10
    C = rand(600,4000);
    d = rand(600);
    @time C\d;
end
```

```
  0.347197 seconds (3.64 k allocations: 41.110 MiB)
  0.342625 seconds (3.64 k allocations: 41.110 MiB)
  0.358289 seconds (3.64 k allocations: 41.110 MiB)
  0.343822 seconds (3.64 k allocations: 41.110 MiB)
  0.348632 seconds (3.64 k allocations: 41.110 MiB)
  0.365883 seconds (3.64 k allocations: 41.110 MiB, 2.04% gc time)
  0.352147 seconds (3.64 k allocations: 41.110 MiB, 1.18% gc time)
  0.353465 seconds (3.64 k allocations: 41.110 MiB, 0.66% gc time)
  0.362876 seconds (3.64 k allocations: 41.110 MiB, 0.64% gc time)
  0.356809 seconds (3.64 k allocations: 41.110 MiB, 0.65% gc time)
```

Figure 1: Result of 16.2.

The complexity of solving the least squares problem with $m \times n$ matrix A is around $2mn^2$ flops. So my approximate flop rate is $5.48 * 10^{10}$ flop/sec or 55 Gflop/sec.

A16.3 (25 pts)

According to the optimality conditions, we have

$$\begin{bmatrix} 2\boldsymbol{A}^T\boldsymbol{A} & \boldsymbol{C}^T \\ \boldsymbol{C} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{x}} \\ \hat{\boldsymbol{z}} \end{bmatrix} = \begin{bmatrix} 2\boldsymbol{A}^T\boldsymbol{b} \\ \boldsymbol{d} \end{bmatrix}$$

So the two lines of Julia code is shown as below

```
xz = [2*A'A C';C zeros(p,p)]\[2*A'b; d]
x = xz[1:n]
```

The following is also correct:

$$\begin{bmatrix} \boldsymbol{A}^\top\boldsymbol{A} & \boldsymbol{C}^\top \\ \boldsymbol{C} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{z} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}^\top\boldsymbol{b} \\ \boldsymbol{d} \end{bmatrix}$$

The corresponding code is:

```
xz = [A'A C';C zeros(p,p)]\[A'b; d]
x = xz[1:n]
```

A16.4 (25 pts)

The colleague is correct.

According to the QR factorization method for solving constrained least squares as in Algorithm 16.2 in the textbook, we get the result

$$R\hat{x} = Q_1^T b - (1/2)Q_2^T w.$$
$$\tilde{R}w = 2\tilde{Q}^T Q_1^T b - 2\tilde{R}^{-T} d.$$

Substituting w in first line by the second equation, we have

$$\hat{x} = R^{-1}Q_1^T b - R^{-1}Q_2^T \tilde{R}^{-1}\tilde{Q}^T Q_1^T b + R^{-1}Q_2^T \tilde{R}^{-1}\tilde{R}^{-T} d$$
$$\hat{x} = (R^{-1}Q_1^T - R^{-1}Q_2^T \tilde{R}^{-1}\tilde{Q}^T Q_1^T)b + (R^{-1}Q_2^T \tilde{R}^{-1}\tilde{R}^{-T})d.$$

So $F = R^{-1}Q_1^T - R^{-1}Q_2^T \tilde{R}^{-1}\tilde{Q}^T Q_1^T$ and $G = R^{-1}Q_2^T \tilde{R}^{-1}\tilde{R}^{-T}$.

4 (25 pts)

Define the residual vector $\boldsymbol{y} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$, then $\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|^2 = \|\boldsymbol{y}\|^2 = \|[\boldsymbol{0} \quad \boldsymbol{I}] \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} \|^2$. Then the unconstrained least square problem is equivalent to

$$\begin{aligned} \underset{(\boldsymbol{x},\boldsymbol{y})}{\text{minimize}} \quad & \|\boldsymbol{y}\|^2 \\ \text{subject to} \quad & \boldsymbol{y} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}. \end{aligned}$$

We write it in the canonical form

$$\begin{aligned} \underset{(\boldsymbol{x},\boldsymbol{y})}{\text{minimize}} \quad & \left\| [\boldsymbol{0} \quad \boldsymbol{I}] \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} \right\|^2, \\ \text{subject to} \quad & [\boldsymbol{A} \quad \boldsymbol{I}] \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} = \boldsymbol{b}. \end{aligned}$$

For this constrained least square problem, the KKT equations is

$$\begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A}^T \\ \boldsymbol{0} & \boldsymbol{I} & \boldsymbol{I} \\ \boldsymbol{A} & \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \\ \boldsymbol{z} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{b} \end{bmatrix}$$

Because

$$\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{0}, \quad \boldsymbol{A}^T\boldsymbol{z} = \boldsymbol{0} \Rightarrow \boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{0}$$

the KKT equations reduce to

$$\begin{bmatrix} \boldsymbol{0} & \boldsymbol{A}^T \\ \boldsymbol{A} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{b} \end{bmatrix},$$

so that the solution can be obtained by the LDLT factorization of a $(m+n) \times (m+n)$ matrix $\begin{bmatrix} \boldsymbol{0} & \boldsymbol{A}^T \\ \boldsymbol{A} & \boldsymbol{I} \end{bmatrix}$.