Problem 0.1. Watkins 2.6.6

Solution. I wrote a short script to run this for me.

```matlab
for i=1:4
    n=4*i
    z=ones(n,1);
    H=hilb(n);
    b=H*z;
    xhat=H\b
    cond(H,2)
    err=norm(xhat-z,2)
    res=norm(b-H*xhat,2)
end
```

I won’t list the vectors, but the condition numbers were $1.6 \times 10^4$, $1.5 \times 10^{10}$, $1.7 \times 10^{16}$ and $4.2 \times 10^{17}$ for $n = 4, 8, 12, 16$ respectively. The errors were $3 \times 10^{-13}$, $9 \times 10^{-7}$, $8.7 \times 10^{-4}$ and 0.11. The residual norms were $1.6 \times 10^{-16}$ up through $1.1 \times 10^{-15}$, i.e., all very small. In other words, the residuals were small, but the error was large because the condition numbers were so large. □

Problem 0.2. Watkins 2.7.12

Solution. One example is $G = \begin{bmatrix} 1 & 2 \\ 10^{-16} & 2 \end{bmatrix}$ and $\delta G = \begin{bmatrix} 5 \times 10^{-16} & 10^{-16} \\ -10^{-16} & 3 \times 10^{-17} \end{bmatrix}$. In this case, $\|\delta G\|_{\infty}/\|G\|_{\infty} = 2 \times 10^{-16}$ while $|\delta g_{21}|/|g_{21}| = 1$. □

Problem 0.3. Watkins 3.1.5

Solution. (a) The system is $\begin{bmatrix} 1 & 1 \\ 1 & 1.5 \\ 1 & 2 \\ 1 & 2.5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} b \\ m \end{bmatrix} = \begin{bmatrix} 1.1 \\ 1.2 \\ 1.3 \\ 1.3 \\ 1.4 \end{bmatrix}$, where we are trying to find the line $y = mt + b$ that best fits the data.

(b) I found $(b, m) = (0.98, 0.14)$, so $y \approx 0.14t + 0.98$.

(c)
(d) \[ ||r||_2 = 0.054772. \]

**Problem 0.4. Watkins 3.1.8**

*Solution.* My system is
\[
\begin{bmatrix}
1 & 2.7 & 0.37 \\
1 & 4.5 & 0.22 \\
1 & 7.4 & 0.14 \\
1 & 12.2 & 0.082 \\
1 & 20.1 & 0.50
\end{bmatrix}
\begin{bmatrix}
a \\
b \\
c
\end{bmatrix}
= \begin{bmatrix}
1.1 \\
1.2 \\
1.3 \\
1.3 \\
1.4
\end{bmatrix},
\]
where I am trying to fit my data to the model \( a + b e^t + c e^{-t} \). (I didn’t actually round in MATLAB.) Solving this, I get \((a, b, c) = (1.325, 0.00483, -0.641)\). Though the question didn’t ask for it, here’s a graph.

**Problem 0.5. Watkins 3.5.1**
Solution. (a) If $x, y \in S^\perp$, and $z \in S$, then

$$\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle = 0 + 0$$

and so $x + y$ is orthogonal to anything in $S$.

(b) If $x \in S^\perp$, $c$ is a number, and $z \in S$, then

$$\langle cx, z \rangle = c \langle x, z \rangle = c \cdot 0 = 0,$$

so $cx$ is orthogonal to anything in $S$. \[ \square \]

Problem 0.6. Watkins 3.5.2

Solution. By definition of a basis, any vector $x \in \mathbb{R}^n$ can be written as

$$x = c_1 q_1 + \cdots + c_n q_n.$$ 

Since they are orthonormal, $\langle x, q_i \rangle = c_i$. If $s \in S$, then $s = s_1 q_1 + \cdots + s_k q_k$, and so, since we can break up inner products over addition, $\langle x, c \rangle$ is zero (for every $s$) if and only if each of $c_1, \cdots c_k$ are zero. (Otherwise that $c_i$ term would show up in the inner product.)

In other words, $x \in S^\perp$ if and only if $c_1, \cdots c_k$ are zero, and so

$$x = c_{k+1} q_{k+1} + \cdots + c_n q_n.$$ 

In other words, if and only if $x$ is in the span of $q_{k+1}, \cdots q_n$. \[ \square \]

Problem 0.7. Watkins 3.5.7

Solution. (a) To show that it is a subspace, we need to show that if $x, y \in N(A)$, and $c$ is a number, then $x + y$ and $cx$ are in $N(A)$. These are easy calculations:

$$A(x + y) = Ax + Ay = 0 + 0 = 0$$

$$A(cx) = cAx = c \cdot 0 = 0$$

(b) If $x, y \in R(A)$, then there are vectors $\hat{x}$ and $\hat{y}$ such that $A \hat{x} = x$ and $A \hat{y} = y$. Then, $x + y$ is in the range of $A$ since $A(\hat{x} + \hat{y}) = x + y$ and $cx$ is in the range of $A$ since $A(c\hat{x}) = cx$. \[ \square \]

Problem 0.8. Watkins 3.5.11

Solution. We want to show that if $x \in N(A^T)$, then $x \in R(A)^\perp$. If $x \in N(A^T)$, then $A^T x = 0$. To show that $x \in R(A)^\perp$, we want to show that $\langle x, y \rangle = 0$ for any $y \in R(A)$. If $y \in R(A)$, then there is some $\hat{y}$ so that $A \hat{y} = y$. We calculate

$$\langle x, y \rangle = \langle x, A \hat{y} \rangle = \langle A^T x, \hat{y} \rangle = \langle 0, \hat{y} \rangle = 0,$$

and so $x \in R(A)^\perp$. \[ \square \]

Problem 0.9. Watkins 3.5.18a
Solution. To prove “(thing A) if and only if (thing B),” you have to prove “If (thing A) then (thing B)” and “If (thing B) then (thing A).” We will do that.

If $A\hat{x} = y$, then we can subtract the two equations to get $Ax - A\hat{x} = y - y$, which can be simplified to $A(x - \hat{x}) = 0$, which exactly says $x - \hat{x} \in N(A)$.

If $x - \hat{x} \in N(A)$, then $A(x - \hat{x}) = 0$, which can be rearranged as $A\hat{x} = Ax = y$. □

Problem 0.10. Watkins 3.5.23

Solution.

\[
\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} 9 \\ 14 \end{bmatrix},
\]

and so $x = 7$. □

Problem 0.11. Explain why using a tiny pivot can be disastrous.

Solution. If you unnecessarily use a tiny pivot, the calculated multipliers will be very large. Using a large multiplier when you subtract off the multiples of the given row means that the rows of your remaining submatrix will be nearly multiples of each other, since the original data there will have been swamped by the large multiple of the given row. Since they are nearly multiples of each other, in the next step of elimination, there will most likely be catastrophic cancelation, leading to large errors. □

Problem 0.12. Explain why we use Gaussian elimination with partial pivoting even though it is not backwards stable.

Solution. The main reason is that while elimination can be unstable (if, say $\|U\| \approx 2^{n-1}$), it is very rare. Most matrices behave well, even if there are a few that behave badly. Plus, if you suspect that your matrix might be one of the few that misbehaves, you can always check, a posteriori, whether your calculation was small, but checking whether the residual and/or $\|U\|\|L\|/\|A\|$ were small.

We could, of course, use Gaussian elimination with complete pivoting, but that is computationally more expensive, so we tend to avoid that. □