Thanks to Gary Wright for writing up some of these solutions.

0. Read sections 3.1, 3.2, 14.1, and 14.2 in the book. (Note: in Assignment # 5, you were asked to read section 12.3. Make sure you heave read that before beginning this homework.)

1. Page 420, #3.3. Let f be a strictly convex quadratic function of one variable. Prove that the secant method for minimization will terminate in exactly one iteration for any initial start points x_0 and x_1. **Solution.** Since f is quadratic and of one variable we can write $f(x) = ax^2 + bx + c$ for some scalars (numbers) a, b, c and $a \neq 0$. Since f is strictly convex we also know that $a > 0$. Next, since we have $f(x)$ we can write $f'(x) = 2ax + b$. To minimize the function, note that $f'(x) = 2ax + b$, so if we set $0 = f'(x_*) = 2ax_* + b$, then $x_* = \frac{-b}{2a}$ where x_* is the minimizer. (Note that it is easy to find the minimizer by basic algebra, so we are doing this exercise mainly to test if the secant method behaves like we expect it to behave.) Recall the secant method in one dimension, given by:

$$x_{k+1} = x_k - \frac{(x_k - x_{k-1})}{f'(x_k) - f'(x_{k-1})} f'(x_k)$$

Using this and $f(x) = ax^2 + bx + c$ in the secant method with $k = 1$, we find:

$$\begin{aligned}
x_2 &= x_1 - \frac{(x_1 - x_0)}{f'(x_1) - f'(x_0)} f'(x_1) \\
&= x_1 - \frac{(x_1 - x_0)}{(2ax_1 + b) - (2ax_0 + b)} (2ax_1 + b) \\
&= x_1 - \frac{(x_1 - x_0)}{2ax_1 - 2ax_0} (2ax_1 + b) \\
&= x_1 - \frac{(x_1 - x_0)}{2a(x_1 - x_0)} (2ax_1 + b) \\
&= x_1 - \frac{1}{2a} (2ax_1 + b) \\
&= x_1 - x_1 - \frac{b}{2a} \\
&= \frac{-b}{2a}.
\end{aligned}$$

Note that exactly the same result would hold if we looked for x_3, x_4, etc. Thus, the sequences is constant after x_1. Moreover, since the first iteration finding x_2 is equal to $x_* = \frac{-b}{2a}$ where x_* is the minimizer, x_2 is the minimizer and the secant method terminated after one iteration.

2. Page 421, #3.7. Let B_{k+1} be obtained from B_k using the update formula

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k)v^T}{v^T s_k}$$

where v is a vector such that $v^T s_k \neq 0$. Prove that $B_{k+1} s_k = y_k$.

Proof.

$$\begin{aligned} B_{k+1} &= B_k + \frac{(y_k - B_k s_k)v^T}{v^T s_k} \\ &= B_k + \frac{(y_k - B_k s_k)}{s_k} \\ &= \frac{B_k s_k + y_k - B_k s_k}{s_k} \\ &= \frac{y_k}{s_k} \\ B_{k+1} s_k &= y_k \end{aligned}$$

□

3. Page 82, #1.3. Consider the system of inequality constraints $Ax \geq b$ with

$$A = \begin{pmatrix} 9 & 4 & 1 & 9 & -7 \\ 6 & -7 & 8 & -4 & -6 \\ 1 & 6 & 3 & -7 & 6 \end{pmatrix} \text{ and } b = \begin{pmatrix} -15 \\ -30 \\ -20 \end{pmatrix}$$

For the given values of x and p, perform a ratio test to determine the maximum step length $\overline{\alpha}$ such that $x + \overline{\alpha} p$ remains feasible.

Solution.
We want to find the a_is that satisfy $a_i^T p < 0$. Then we can use the ratio test to find the smallest of such a_is to determine the value of $\overline{\alpha}$. Let

$$\begin{aligned} a_1^T &= (9, 4, 1, 9, -7) \text{ and } b_1 = -15 \\ a_2^T &= (6, -7, 8, -4, -6) \text{ and } b_2 = -30 \\ a_3^T &= (1, 6, 3, -7, 6) \text{ and } b_3 = -20 \end{aligned}$$

(i) $x = (8, 4, -3, 4, 1)^T$ and $p = (1, 1, 1, 1, 1)^T$

$$a_1^T p = 16 \not< 0 \qquad a_2^T p = -3 < 0 \qquad a_3^T p = 9 \not< 0$$

$$\overline{\alpha} = \min\{\frac{a_2^T x - b_2}{(-a_2^T p)}\} = \min\{\frac{-26-(-30)}{-(-3)}\} = \frac{4}{3}$$

So the maximum step length will occur when $\overline{\alpha} = \frac{4}{3}$.

(ii) $x = (7,-4,-3,-3,3)^T)$ and $p = (3,2,0,1,-2)^T$

$$a_1^T p = 58 \not< 0 \qquad a_2^T p = 12 \not< 0 \qquad a_3^T p = -4 < 0$$

$$\overline{\alpha} = \min\{\frac{a_3^T x - b_3}{(-a_3^T p)}\} = \min\{\frac{13-(-20)}{-(-4)}\} = \frac{33}{4}$$

So the maximum step length will occur when $\overline{\alpha} = \frac{33}{4}$.

(iii) $x = (5,0,-6,-8,-3)^T)$ and $p = (5,0,5,1,3)^T$

$$a_1^T p = 38 \not< 0 \qquad a_2^T p = 48 \not< 0 \qquad a_3^T p = 31 \not< 0$$

Since $a_i^T p \geq 0$ for all a_i, the constraint will remain satisfied for any $\alpha \geq 0$.

(iv) $x = (9,1,-1,6,3)^T)$ and $p = (-4,-2,4,-2,2)^T$

$$a_1^T p = -72 < 0 \qquad a_2^T p = 18 \not< 0 \qquad a_3^T p = 22 \not< 0$$

$$\overline{\alpha} = \min\{\frac{a_1^T x - b_1}{(-a_1^T p)}\} = \min\{\frac{117-(-15)}{-(-72)}\} = \frac{11}{6}$$

So the maximum step length will occur when $\overline{\alpha} = \frac{11}{6}$.

4. Page 84-85, #2.1 (i) and (iii). In each of the following cases, compute a basis matrix for the null space of the matrix and express the points x_i as $x_i = p_i + q_i$ where p_i is in the null space of A and q_i is in the range space of A^T.

Solution.

(i) $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$, $x_1 = \begin{pmatrix} 1 \\ 3 \\ 1 \\ 2 \end{pmatrix}$, $x_2 = \begin{pmatrix} 0 \\ -2 \\ -3 \\ 4 \end{pmatrix}$

$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix} \sim A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \Rightarrow$ basis for $\text{Nul}(A) = Z =$

$$\begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$$

For this problem we'll find $\vec{p}$ via the equation $\vec{p} = (I - A^T(AA^T)^{-1}A)\vec{x}$ and then $\vec{q}$ via $\vec{q} = \vec{x} - \vec{p}$.

$$A^T = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 1 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

$$AA^T = \begin{pmatrix} 4 & 0 & 2 \\ 0 & 4 & 0 \\ 2 & 0 & 2 \end{pmatrix}$$

$$(AA^T)^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & \frac{1}{4} & 0 \\ -\frac{1}{2} & 0 & 1 \end{pmatrix}$$

$$A^T(AA^T)^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & -\frac{1}{2} \\ 0 & -\frac{1}{4} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{4} & -\frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

$$A^T(AA^T)^{-1}A = \begin{pmatrix} \frac{3}{4} & -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ -\frac{1}{4} & \frac{3}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{3}{4} & -\frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

$$I - (A^T(AA^T)^{-1}A) = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

$$\vec{p}_1 = I - (A^T(AA^T)^{-1}A)\vec{x}_1 = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{4} \\ -\frac{1}{4} \\ -\frac{1}{4} \end{pmatrix} \text{ and thus, } \vec{q}_1 = \vec{x}_1 - \vec{p}_1 = \begin{pmatrix} 1 \\ 3 \\ 1 \\ 2 \end{pmatrix} - \begin{pmatrix} \frac{1}{4} \\ \frac{1}{4} \\ -\frac{1}{4} \\ -\frac{1}{4} \end{pmatrix} = \begin{pmatrix} \frac{3}{4} \\ \frac{11}{4} \\ \frac{5}{4} \\ \frac{9}{4} \end{pmatrix}$$

$$\vec{p}_2 = I - (A^T(AA^T)^{-1}A)\vec{x}_2 = \begin{pmatrix} -\frac{3}{4} \\ -\frac{3}{4} \\ \frac{3}{4} \\ \frac{3}{4} \end{pmatrix}$$

and thus,

$$\vec{q}_2 = \vec{x}_2 - \vec{p}_2 = \begin{pmatrix} 0 \\ -2 \\ -3 \\ 4 \end{pmatrix} - \begin{pmatrix} -\frac{3}{4} \\ -\frac{3}{4} \\ \frac{3}{4} \\ \frac{3}{4} \end{pmatrix} = \begin{pmatrix} \frac{3}{4} \\ -\frac{5}{4} \\ -\frac{15}{4} \\ \frac{13}{4} \end{pmatrix}$$

(iii) $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$, $x_1 = \begin{pmatrix} 4 \\ 3 \\ 4 \\ 0 \end{pmatrix}$, $x_2 = \begin{pmatrix} -1 \\ 1 \\ 5 \\ -5 \end{pmatrix}$

$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \sim A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \Rightarrow$

basis for $\text{Nul}(A) = Z = \begin{pmatrix} 0 & -1 \\ -1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$

For this method we'l find $\vec{p}$ by orthogonalizing the columns of Z and projecting the basis. Let $\hat{z}_i$ be the i^{th} column vector of Z.

$$\vec{p}_1 = \frac{\vec{x}_1 \cdot \hat{z}_1}{\|\hat{z}_1\|^2}\hat{z}_1 + \frac{\vec{x}_1 \cdot \hat{z}_2}{\|\hat{z}_2\|^2}\hat{z}_2 = \frac{1}{2}\hat{z}_1 + -\frac{4}{2}\hat{z}_2 = \begin{pmatrix} 0 \\ -\frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix} + \begin{pmatrix} 2 \\ 0 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ -\frac{1}{2} \\ \frac{1}{2} \\ -2 \end{pmatrix} \quad \text{and thus,}$$

$$\vec{q}_1 = \vec{x}_1 - \vec{p}_1 = \begin{pmatrix} 4 \\ 3 \\ 4 \\ 0 \end{pmatrix} - \begin{pmatrix} 2 \\ -\frac{1}{2} \\ \frac{1}{2} \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ \frac{7}{2} \\ \frac{7}{2} \\ 2 \end{pmatrix}$$

$$\vec{p}_2 = \frac{\vec{x}_2 \cdot \hat{z}_1}{\|\hat{z}_1\|^2}\hat{z}_1 + \frac{\vec{x}_2 \cdot \hat{z}_2}{\|\hat{z}_2\|^2}\hat{z}_2 = \frac{4}{2}\hat{z}_1 + -\frac{4}{2}\hat{z}_2 = \begin{pmatrix} 0 \\ -2 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 2 \\ 0 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 2 \\ -2 \end{pmatrix} \quad \text{and thus,}$$

$$\vec{q}_2 = \vec{x}_2 - \vec{p}_2 = \begin{pmatrix} -1 \\ 1 \\ 5 \\ -5 \end{pmatrix} - \begin{pmatrix} 2 \\ -2 \\ 2 \\ -2 \end{pmatrix} = \vec{q}_2 = \begin{pmatrix} -3 \\ 3 \\ 3 \\ -3 \end{pmatrix}$$

5. Page 85, #2.6. Suppose that you are given a matrix A and a vector p and are told that p is in the null space of A. On a computer, you cannot expect that Ap will be exactly equal to zero because of rounding errors. How large would the computed value of $\|Ap\|$ have to be before you could conclude that p was not in the null space of A? If the computed value of $\|Ap\|$ is zero, can you conclude that p is in the null space of A? (Note: This probably is somewhat more open-ended.)

   Solution. Answers may vary widely. The key is to remember that in a computer, when numbers get very small (very close to zero), they are indistinguishable from zero. Try some test matrices in Matlab to see if you can come up with some ideas!