Homework 4

12a. From $\nabla f(\boldsymbol{x}) = \boldsymbol{b} + A\boldsymbol{x}$ we see that $\nabla f(\boldsymbol{x}) = \boldsymbol{0}$ has a unique solution $\boldsymbol{x}^* = -A^{-1}\boldsymbol{b}$. It means that $f(\boldsymbol{x})$ has a unique critical point, which implies, by Theorem 1.2.3, that $f(\boldsymbol{x})$ has at most one global minimizer.

Since $Hf(\boldsymbol{x}) = A$ is positive definite, we deduce from Theorem 1.2.9(b) that $\boldsymbol{x}^*$ is a global minimizer of $f(\boldsymbol{x})$. Since $f(\boldsymbol{x})$ has at most one global minimizer, $\boldsymbol{x}^*$ is its unique global minimizer.

12b. Suppose $\boldsymbol{x}^{(0)} - \boldsymbol{x}^*$ is an eigenvector of A. Then $\boldsymbol{x}^{(0)} - \boldsymbol{x}^* \neq \boldsymbol{0}$ and $A(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*) = \lambda(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*)$, for a number λ. Notice that

$$\begin{aligned}\nabla f(\boldsymbol{x}^{(0)}) &= \boldsymbol{b} + A\boldsymbol{x}^{(0)} = A(A^{-1}\boldsymbol{b} + \boldsymbol{x}^{(0)})\\ &= A(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*) = \lambda(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*).\end{aligned}$$

Thus $\varphi_0(t) = f(\boldsymbol{x}^{(0)} - t\nabla f(\boldsymbol{x}^{(0)})) = f(\boldsymbol{x}^{(0)} - t\lambda(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*))$. Now

$$\begin{aligned}\varphi_0'(t) &= \nabla f(\boldsymbol{x}^{(0)} - t\lambda(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*)) \cdot (-\lambda(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*))\\ &= -\lambda(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*) \cdot \nabla f(\boldsymbol{x}^{(0)} - t\lambda(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*))\\ &= -\lambda(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*) \cdot (\boldsymbol{b} + A(\boldsymbol{x}^{(0)} - t\lambda(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*)))\\ &= -\lambda(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*) \cdot A(A^{-1}\boldsymbol{b} + \boldsymbol{x}^{(0)} - t\lambda(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*))\\ &= -\lambda(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*) \cdot A(-\boldsymbol{x}^* + \boldsymbol{x}^{(0)} - t\lambda(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*))\\ &= -\lambda(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*) \cdot A(1 - t\lambda)(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*)\\ &= -\lambda(1 - t\lambda)(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*)A(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*).\end{aligned}$$

Notice that $(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*)A(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*) > 0$, since A is positive definite and $\boldsymbol{x}^{(0)} - \boldsymbol{x}^* \neq \boldsymbol{0}$. For the same reason, $\lambda \neq 0$ (otherwise we would have $A(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*) = \lambda(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*) = \boldsymbol{0}$, so $\boldsymbol{x}^{(0)} - \boldsymbol{x}^* = A^{-1}\boldsymbol{0} = \boldsymbol{0}$, which is not the case). Thus $\varphi_0'(t) = 0$ has a unique solution $t^* = 1/\lambda$. Since $\varphi_0''(t) = \lambda^2(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*)A(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*) > 0$, we deduce from Theorem 1.2.9(b) that t^* is a global minimizer of $\varphi_0(t)$, and thus $t_0 = 1/\lambda$. Therefore,

$$\begin{aligned}\boldsymbol{x}^{(1)} &= \boldsymbol{x}^{(0)} - t_0\nabla f(\boldsymbol{x}^{(0)})\\ &= \boldsymbol{x}^{(0)} - t_0\lambda(\boldsymbol{x}^{(0)} - \boldsymbol{x}^*) = \boldsymbol{x}^{(0)} - (\boldsymbol{x}^{(0)} - \boldsymbol{x}^*) = \boldsymbol{x}^*,\end{aligned}$$

which is what we want to prove.

14a. Let $g(\boldsymbol{x}) = \nabla f(\boldsymbol{x}) = (4x_1 - x_2, 2x_2 - x_1)$. Then

$$\begin{aligned}\boldsymbol{x}^{(1)} &= \boldsymbol{x}^{(0)} - D_0^{-1}g(\boldsymbol{x}^{(0)}) = (1,4) - g(1,4) = (1,-3).\\ \boldsymbol{d}^{(0)} &= \boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)} = (1,-3) - (1,4) = (0,-7);\\ \boldsymbol{y}^{(0)} &= g(\boldsymbol{x}^{(1)}) - g(\boldsymbol{x}^{(0)}) = g(1,-3) - g(1,4) = (7,-14).\end{aligned}$$

$$(\boldsymbol{y}^{(0)} - D_0\boldsymbol{d}^{(0)}) \otimes \boldsymbol{d}^{(0)} = (7,-7) \otimes (0,-7) = \begin{bmatrix}0 & -49\\ 0 & 49\end{bmatrix},$$

$$D_1 = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \frac{1}{49}\begin{bmatrix}0 & -49\\ 0 & 49\end{bmatrix} = \begin{bmatrix}1 & -1\\ 0 & 2\end{bmatrix}.$$

$$\begin{aligned}\boldsymbol{x}^{(2)} &= \boldsymbol{x}^{(1)} - D_1^{-1}g(\boldsymbol{x}^{(1)}) = (1,-3) - D_1^{-1}g(1,-3)\\ &= \begin{bmatrix}1\\ -3\end{bmatrix} - \begin{bmatrix}1 & \frac{1}{2}\\ 0 & \frac{1}{2}\end{bmatrix}\begin{bmatrix}7\\ -7\end{bmatrix} = \begin{bmatrix}-\frac{5}{2}\\ \frac{1}{2}\end{bmatrix}.\end{aligned}$$

14b. $D_0 = Hf(1,4) = \begin{bmatrix}4 & -1\\ -1 & 2\end{bmatrix}$ and $D_0^{-1} = \frac{1}{7}\begin{bmatrix}2 & 1\\ 1 & 4\end{bmatrix}$. Then

$$\begin{aligned}\boldsymbol{x}^{(1)} &= \boldsymbol{x}^{(0)} - D_0^{-1}g(\boldsymbol{x}^{(0)}) = (1,4) - (1,4) = (0,0).\\ \boldsymbol{d}^{(0)} &= \boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)} = (0,0) - (1,4) = (-1,-4);\\ \boldsymbol{y}^{(0)} &= g(\boldsymbol{x}^{(1)}) - g(\boldsymbol{x}^{(0)}) = g(0,0) - g(1,4) = (0,-7).\\ (\boldsymbol{y}^{(0)} &- D_0\boldsymbol{d}^{(0)}) \otimes \boldsymbol{d}^{(0)} = (0,0) \otimes (-1,-4) = \boldsymbol{0}_{2\times 2},\\ D_1 &= D_0 + \boldsymbol{0} = D_0.\\ \boldsymbol{x}^{(2)} &= \boldsymbol{x}^{(1)} - D_1^{-1}g(\boldsymbol{x}^{(1)}) = (0,0) - D_1^{-1}g(0,0) = (0,0).\end{aligned}$$

In fact, it is easy to check that $(0,0)$ is the unique global minimizer of $f(x_1, x_2)$. This solution is achieved in the first iteration.

20. First, $\nabla f(\boldsymbol{x}) = (2x_1 - x_2, -x_1 + 3x_2)$. With $\boldsymbol{x}^{(0)} = (1,2)$ and $D_0 = I$, we have

$$\varphi_0(t) = f(\boldsymbol{x}^{(0)} - tI^{-1}\nabla f(\boldsymbol{x}^{(0)})) = f(1, 2-5t) = \tfrac{75}{2}t^2 - 25t + 5,$$

which has a unique minimizer $t_0 = 1/3$. It follows that

$$\begin{aligned}\boldsymbol{x}^{(1)} &= (1, 2 - 5\cdot\tfrac{1}{3}) = (1, \tfrac{1}{3}),\\ \boldsymbol{d}^{(0)} &= \boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)} = (0, -\tfrac{5}{3}),\\ \boldsymbol{y}^{(0)} &= \nabla f(\boldsymbol{x}^{(1)}) - \nabla f(\boldsymbol{x}^{(0)}) = (\tfrac{5}{3}, -5),\end{aligned}$$

$$\begin{aligned}D_1 &= D_0 + \frac{\boldsymbol{y}^{(0)} \otimes \boldsymbol{y}^{(0)}}{\boldsymbol{y}^{(0)} \cdot \boldsymbol{d}^{(0)}} - \frac{D_0\boldsymbol{d}^{(0)} \otimes D_0\boldsymbol{d}^{(0)}}{\boldsymbol{d}^{(0)}D_0\boldsymbol{d}^{(0)}}\\ &= I + \frac{3}{25}\begin{bmatrix}\frac{25}{9} & -\frac{25}{3}\\ -\frac{25}{3} & 25\end{bmatrix} - \frac{9}{25}\begin{bmatrix}0 & 0\\ 0 & \frac{25}{9}\end{bmatrix} = \begin{bmatrix}\frac{4}{3} & -1\\ -1 & 3\end{bmatrix}.\end{aligned}$$

$$D_1^{-1}\nabla f(\boldsymbol{x}^{(1)}) = \begin{bmatrix}1 & \frac{1}{3}\\ \frac{1}{3} & \frac{4}{9}\end{bmatrix}\begin{bmatrix}\frac{5}{3}\\ 0\end{bmatrix} = \begin{bmatrix}\frac{5}{3}\\ \frac{5}{9}\end{bmatrix},$$

$$\begin{aligned}\varphi_1(t) &= f(\boldsymbol{x}^{(1)} - tD_1^{-1}\nabla f(\boldsymbol{x}^{(1)})) = f(1 - \tfrac{5t}{3}, \tfrac{1}{3} - \tfrac{5t}{9}),\\ \varphi_1'(t) &= \nabla f(1 - \tfrac{5t}{3}, \tfrac{1}{3} - \tfrac{5t}{9}).(-\tfrac{5}{3}, -\tfrac{5}{9}) = \tfrac{25}{27}(5t - 3),\end{aligned}$$

so $t_1 = \frac{3}{5}$ and $\boldsymbol{x}^{(2)} = (1 - \frac{5t_1}{3}, \frac{1}{3} - \frac{5t_1}{9}) = (0,0)$.

1. In matrix form we have

$$A = \begin{bmatrix}1 & 1 & 1\\ 0 & 0 & 1\\ 1 & 0 & 1\\ 2 & 0 & 5\\ -7 & 8 & 0\\ 1 & 2 & -1\end{bmatrix} \quad\text{and}\quad b = \begin{bmatrix}3\\ 1\\ 2\\ 8\\ 0\\ 1\end{bmatrix}. \quad\text{Then}$$

$$A^tA = \begin{bmatrix}56 & -53 & 11\\ -53 & 69 & -1\\ 11 & -1 & 29\end{bmatrix} \quad\text{and}$$

$$(A^tA)^{-1} = \tfrac{1}{23356}\begin{bmatrix}2000 & 1526 & -706\\ 1526 & 1503 & -527\\ -706 & -527 & 1055\end{bmatrix} \approx \begin{bmatrix}.086 & .065 & -.03\\ .065 & .064 & -.02\\ -.03 & -.02 & .045\end{bmatrix}$$

and $(A^tA)^{-1}A^t$

$$\begin{aligned}&= \tfrac{1}{23356}\begin{bmatrix}2820 & -706 & 1294 & 470 & -1792 & 5758\\ 2502 & -527 & 999 & 417 & 1342 & 5059\\ -178 & 1055 & 349 & 3863 & 726 & -2815\end{bmatrix}\\ &\approx \begin{bmatrix}.12 & -.03 & .06 & .02 & -.08 & .25\\ .11 & -.02 & .04 & .02 & .06 & .22\\ -.01 & .05 & .01 & .17 & .03 & -.12\end{bmatrix}\end{aligned}$$

and thus the least squares solution is

$$(A^tA)^{-1}A^tb = \tfrac{1}{5839}(4965, 4343, 7327) \approx (0.85, 0.74, 1.25).$$