1. Consider the Rosenbrock function:

$$f(\mathbf{x}) = 100(y - x^2)^2 + (1 - x)^2$$

   (a) Compute the the gradient and Hessian of f. Is f a convex function?

   Solution. $\nabla f(x) = [-400x^3 - 400xy + 2x - 2, 200(y - x^2)]^T$. From this, we compute

$$H(x, y) = \begin{bmatrix} 400(3x^2 - y) + 2 & -400x \\ -400x & 200 \end{bmatrix}$$

   For $y = 1, x = 0$, the first leading principal minor is negative, so the Hessian is not positive semidefinite for all values of $\mathbf{x}$. Thus, $f(\mathbf{x})$ is not a convex function.

   (b) Find all stationary points of f (if there are more than one). Are they global minimizers of f?

   Solution. The stationary points are solutions to $\nabla f(\mathbf{x}) = \mathbf{0}$. The second constraint requires $y = x^2$. Substituting into the first equation, we get that $-400x^3 - 400x^3 + 2x - 2 = 0 \implies x = 1$. So $(x, y) = (1, 1)$ is our only stationary point. Since $\det H(1, 1) = 2$ and $H(1, 1)_{1,1} = 802$, our matrix is positive definite at $(1, 1)$. Thus, it is a local minimizer. f is not a convex function, so this may not necessarily be a global minimizer.

2. Consider the problem of least squares:

$$\min_{\mathbf{x} \in \mathbb{R}^n} \quad \|A\mathbf{x} - \mathbf{b}\|_2^2$$

   where A is an $m \times n$ matrix and $\mathbf{b} \in \mathbb{R}^m$.

   (a) Write a necessary condition for optimality. Is it also a sufficient condition?

   Solution. A necessary condition for optimality is $\nabla f(x) = 0$ and $\nabla^2 f(x)$ is positive semidefinite. Note here that $\nabla f(x) = A^T(Ax - b)$ and $\nabla^2 f(x) = A^T A$. Since $x^T A^T Ax = \|Ax\|_2^2 \geq 0$, we are always positive semidefinite. This is also a sufficient condition because this is a quadratic function. Otherwise, we would need $x^T A^T Ax = \|Ax\|_2^2 \neq 0 \implies Ax \neq 0$. If A is full column rank, then this is true for nonzero x anyways.

   (b) Is the optimal solution unique? Why or why not?

   Solution. Since we are positive semidefinite and not positive definite, this problem is not strictly convex. Thus, there can be multiple optimal solutions.

   (c) Can you give a closed-form solution of the optimal solution? Specify any assumptions that you may need.

   Solution. From above, we require that $\nabla f(x) = A^T(Ax - b) = 0$. Then $x = (A^T A)^{-1} A^T b$ is an optimal solution provided that $A^T A$ is invertible.

(d) Solve the problem for

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 2 & 2 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 6 \\ 2 \\ 0 \end{bmatrix}.$$

Solution. Solving this via the normal equations, we first compute $A^T A = \begin{bmatrix} 5 & -2 & 1 \\ -2 & 6 & 4 \\ 1 & 4 & 5 \end{bmatrix}$ and $y = A^T b = [4, 12, 12]^T$. Using MATLAB's backslash operator we obtain, $x^* = A^T A \backslash y \approx \begin{bmatrix} 2 & 2.857 & -0.285 \end{bmatrix}^T$.