Seonho Woo (clairewo@umich.edu)

Problem 1.
Gradient and Hessian:

Obj. func.

$$f(x) = x_1 + x_2$$
$$\nabla f(x) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$
$$\nabla^2 f(x) = \mathbf{0}_{2\times 2}$$

Constr.

$$c(x) = x_1^2 + x_2^2 - 2$$
$$\nabla c(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}$$
$$\nabla^2 c(x) = 2I_{2\times 2}$$

Problem 2.
Gradient and Hessian:

Obj. func.

Set *p(x) = x1x2x3x4x5* in the given problem:

$$f(x) = \exp(p(x)) - \frac{1}{2}(x_1^3 + x_2^3 + 1)^2$$

$$\nabla f(x) = \exp(p(x)) \begin{bmatrix} x_2x_3x_4x_5 \\ x_1x_3x_4x_5 \\ x_1x_2x_4x_5 \\ x_1x_2x_3x_5 \\ x_1x_2x_3x_4 \end{bmatrix} - 3(x_1^3 + x_2^3 + 1) \begin{bmatrix} x_1^2 \\ x_2^2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\nabla^2 f(x)(:, 1:3) = \exp(p(x)) \begin{bmatrix} (x_2x_3x_4x_5)^2 & x_3x_4x_5(1+p(x)) & x_2x_4x_5(1+p(x)) \\ x_3x_4x_5(1+p(x)) & (x_1x_3x_4x_5)^2 & x_1x_4x_5(1+p(x)) \\ x_2x_4x_5(1+p(x)) & x_1x_4x_5(1+p(x)) & (x_1x_2x_4x_5)^2 \\ x_2x_3x_5(1+p(x)) & x_1x_3x_5(1+p(x)) & x_1x_2x_5(1+p(x)) \\ x_2x_3x_4(1+p(x)) & x_1x_3x_4(1+p(x)) & x_1x_2x_4(1+p(x)) \end{bmatrix}$$

$$\nabla^2 f(x)(:, 4:5) = \exp(p(x)) \begin{bmatrix} x_2x_3x_5(1+p(x)) & x_2x_3x_4(1+p(x)) \\ x_1x_3x_5(1+p(x)) & x_1x_3x_4(1+p(x)) \\ x_1x_2x_5(1+p(x)) & x_1x_2x_4(1+p(x)) \\ (x_1x_2x_3x_5)^2 & x_1x_2x_3(1+p(x)) \\ x_1x_2x_3(1+p(x)) & (x_1x_2x_3x_4)^2 \end{bmatrix}$$

$$c(x) = \begin{bmatrix} ||x||_2^2 - 10 \\ x_2x_3 - 5x_4x_5 \\ x_1^3 + x_2^3 + 1 \end{bmatrix}$$

$$\nabla c(x) = \begin{bmatrix} 2x_1 & 2x_2 & 2x_3 & 2x_4 & 2x_5 \\ 0 & x_3 & x_2 & -5x_5 & -5x_4 \\ 3x_1^2 & 3x_2^2 & 0 & 0 & 0 \end{bmatrix}$$

$$\nabla^2 c_1(x) = 2I_{5\times 5}$$

$$\nabla^2 c_2(x) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & -5 & 0 \end{bmatrix}$$

$$\nabla^2 c_3(x) = \begin{bmatrix} 6x_1 & 0 & 0 & 0 & 0 \\ 0 & 6x_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

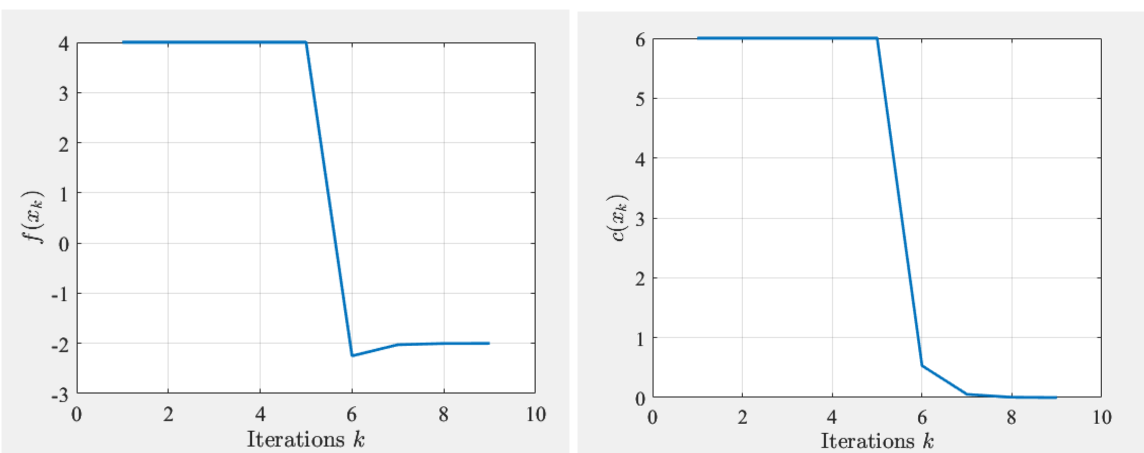
Constr:

==

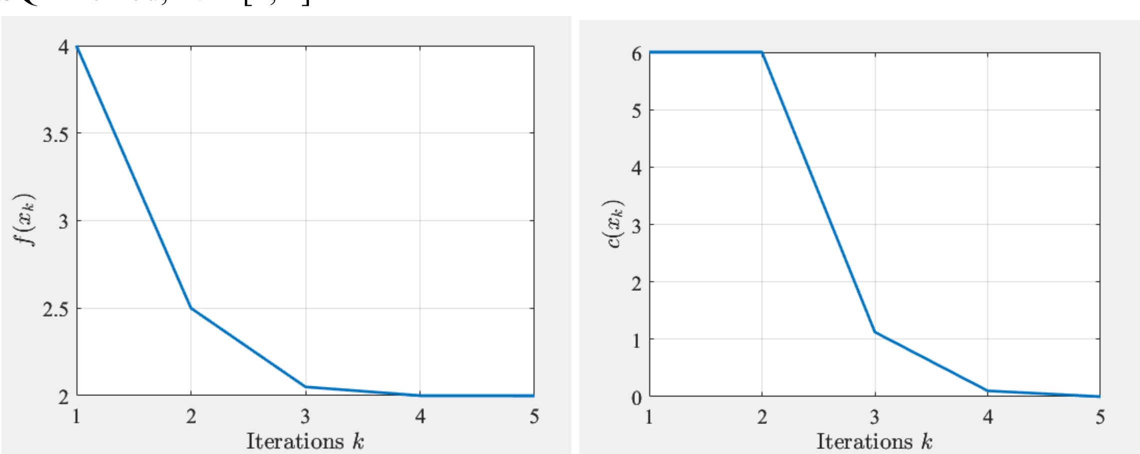
Plots

Problem 1.

QRP method, x0 = [2, 2]T

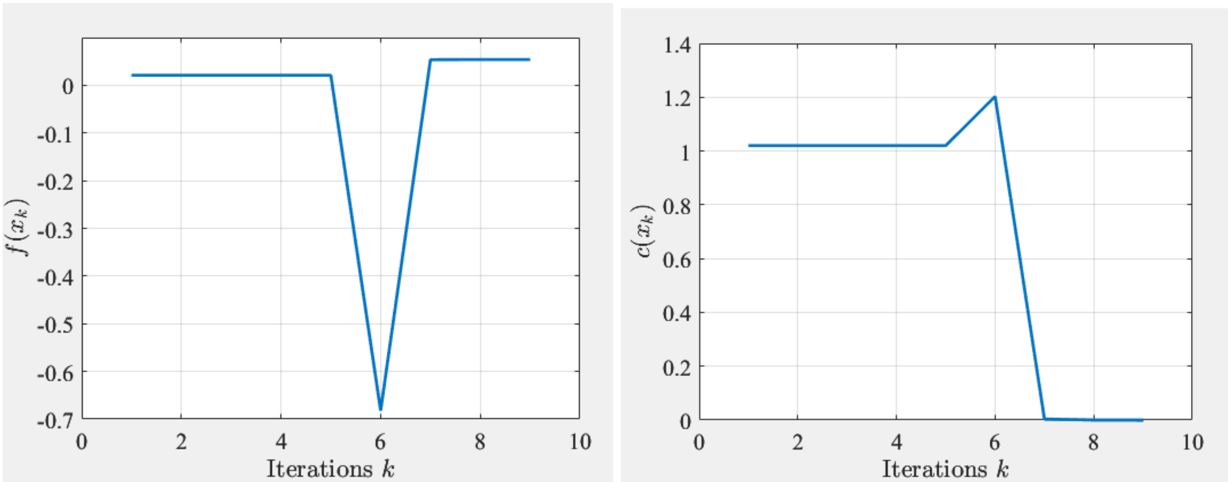


SQP method, x0 = [2, 2]T

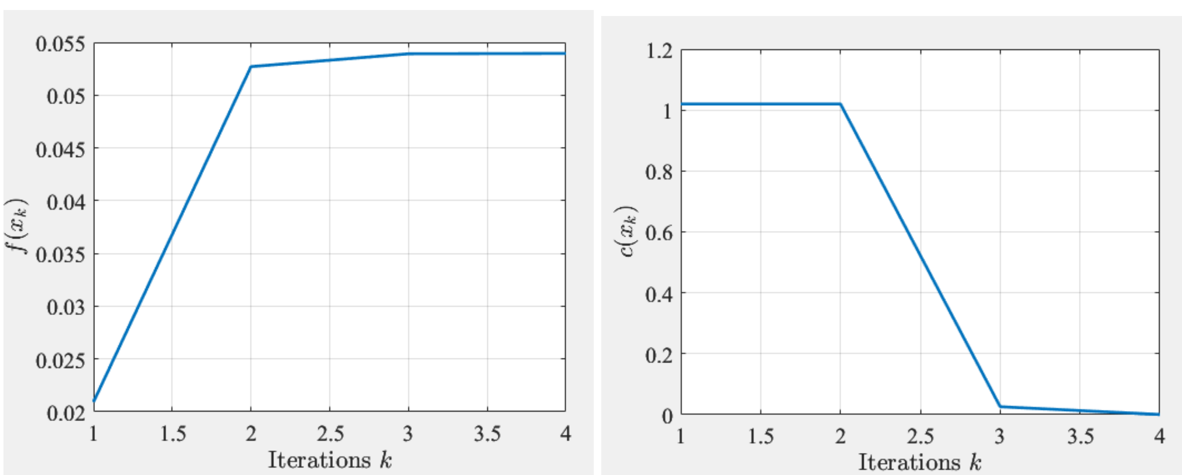


Problem 2.

QRP method, x0 = [-1.8, 1.7, 1.9, -0.8, -0.8]T



SQP method, x0 = [-1.8, 1.7, 1.9, -0.8, -0.8]T



As shown in convergence plots above, all optimization methods achieve convergence for both sets of constrained problems. On the first problem, SQP converges to the maximum value of x* = (1,1)T because the assumption of positive definiteness for Hk does not hold, due to the Hessian being a zero matrix. However, the quadratic penalty method converges to the minimum value for problem 1. For problem 2, both methods converge to the minimum value.

To be specific, the speed at which the quadratic penalty method (QPM) converges varies depending on the optimization problem being tackled and is influenced by factors such as the form of the penalty function, the properties of the objective function, and the constraints. The convergence rate for Problem 1 starts as sublinear and eventually becomes linear, while for

Problem 2, the method initially prioritizes reducing the function value at the expense of meeting the constraints before increasing the function value towards the optimum while satisfying the constraints, resulting in a convergence rate faster than linear.

The convergence rate of the Sequential Quadratic Method (SQM) depends on factors such as the properties of the objective function, the constraints, and the initial solution. Generally, the SQM exhibits superlinear convergence, which is faster than sublinear but slower than linear convergence. For Problem 1, the convergence rate appears slower than linear but still somewhat linear. For Problem 2, the method exhibits superlinear convergence but experiences convergence stagnation after a few iterations, especially when dealing with ill-conditioned or highly nonlinear constraints.

Computational cost:
Generally speaking, the computational cost of SQP is higher than QPM.

QPM involves solving a sequence of unconstrained optimization problems, where each problem has a quadratic penalty term added to the objective function that penalizes violations of the constraints. The computational cost of QPM depends on the number of penalty parameters and the number of iterations required to achieve convergence. The number of penalty parameters needed to achieve convergence increases with the complexity of the problem, which can increase the computational cost.

In contrast, SQP involves solving a sequence of constrained quadratic programming subproblems. Each subproblem involves approximating the objective and constraint functions by a quadratic function and solving the resulting quadratic program. The computational cost of SQP depends on the number of iterations required to achieve convergence, the size of the subproblems, and the complexity of the problem.

While both methods may require several iterations to converge to a solution, the computational cost of SQP can be higher than QPM because it involves solving more complex subproblems at each iteration. However, SQP is known to converge faster than QPM for many problems, which can offset the higher computational cost.