# Exercises for Part 1.

1. Let $c_{\mathcal{E}}(x)$ and $c_{\mathcal{I}}(x)$ be twice-continuously differentiable vector functions of $x$, and let $f \in C^2$. Consider the problem

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \;\; f(x) \;\; \text{subject to both} \;\; c_{\mathcal{E}}(x) = 0 \;\; \text{and} \;\; c_{\mathcal{I}}(x) \geq 0.$$

   (a) Write down first-order necessary optimality conditions for a point $x_*$ to solve this problem (Hint: combine the conditions in Theorems 1.6 and 1.8, using different vectors $y_{\mathcal{E}}$ and $y_{\mathcal{I}}$ for the equality and inequality constraints).

   (b) Write down second-order necessary optimality conditions for a point $x_*$ to solve this problem (Hint: combine the conditions in Theorems 1.7 and 1.9).

2. Suppose that $f_i(x)$, $i = 1, \ldots, m$, are twice-continuously differentiable functions of $x$. Consider the *non-differentiable* optimization problem

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \;\; f(x) = \max_{1 \leq i \leq m} |f_i(x)|. \tag{1}$$

   (a) Why might this problem be "non-differentiable"?

   (b) Argue that this problem is equivalent to the *differentiable* problem

$$\underset{x \in \mathbb{R}^n u \in \mathbb{R}}{\text{minimize}} \;\; u \;\; \text{subject to} \;\; -u \leq f_i(x) \leq u$$

   for some additional variable $u$.

   (c) Hence or otherwise deduce first-order necessary optimality conditions for (1).

# Exercises for Part 2.

1. Consider applying the method of steepest descent with exact line–searches to the problem

$$\underset{x}{\text{minimize}} \; \tfrac{1}{2} x^T H x,$$

where $H$ is a positive definite Hessian.

(a) Show that $x_* = 0$ solves the above problem.

(b) Show that the steplength $\alpha$ obtained by performing an exact line–search from $x$ in the direction $p$ is given by

$$\alpha = -\frac{p^T g}{p^T H p},$$

where $g$ is the gradient at $x$.

(c) Let $H$ be a diagonal matrix given by

$$H = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}, \text{ where } \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0.$$

If the starting point $x_1 = \left( \frac{\sigma}{\lambda_1}, 0, \ldots, 0, \frac{1}{\lambda_n} \right)^T$ is chosen, where $\sigma = \pm 1$, show that

$$x_2 = \frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n} \left( \frac{-\sigma}{\lambda_1}, 0, \ldots, 0, \frac{1}{\lambda_n} \right)^T.$$

Hence show that at iteration $k+1$ the iterate is

$$x_{k+1} = \left( \frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n} \right)^k \left( \frac{(-1)^k \sigma}{\lambda_1}, 0, \ldots, 0, \frac{1}{\lambda_n} \right)^T.$$

What can you say about the speed of convergence, if
*(i)* $\lambda_1 = \lambda_n$ and if
*(ii)* $\lambda_1$ is much greater than $\lambda_n$ ($\lambda_1 \gg \lambda_n$) ?

# Exercises for Part 3.

1. Solve the "trust-region" sub-problem

$$\underset{s \in \mathbb{R}^n}{\text{minimize}} \;\; s^T g + \tfrac{1}{2} s^T B s \;\; \text{subject to} \;\; \|s\|_2 \leq \Delta$$

in the following cases:

(a)

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad g = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \;\; \text{and} \;\; \Delta = 2,$$

(b)

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad g = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \;\; \text{and} \;\; \Delta = 5/12$$

[Hint: a root of the nonlinear equation

$$\frac{1}{(1+\lambda)^2} + \frac{1}{(2+\lambda)^2} = 25/144$$

is $\lambda = 2$.],

(c)

$$B = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad g = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \;\; \text{and} \;\; \Delta = 5/12,$$

(d)

$$B = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad g = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \;\; \text{and} \;\; \Delta = 1/2, \;\; \text{and}$$

(e)

$$B = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad g = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \;\; \text{and} \;\; \Delta = \sqrt{2}.$$

# Exercises for Part 4.

1. Consider the problem

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \;\; \tfrac{1}{2}x_1^2 + x_2 \;\; \text{subject to} \;\; x_2 \geq 0$$

   (a) What is the minimizer of this problem? What is the value of its Lagrange multiplier?

   (b) Write down the logarithmic barrier function $\Phi(x, \mu)$ for the problem. What is the minimizer $x(\mu)$ of the barrier function as a function of the barrier parameter $\mu$? What Lagrange multiplier estimate does this minimizer give?

   (c) Compute the Hessian matrix of the logarithmic barrier function. What are its eigenvalues at the minimizer of the barrier function? How do these eigenvalues behave as the barrier parameter decreases to zero?

   (d) Find the primal-dual step at $x(\mu)$ when $\mu$ is reduced to $\bar{\mu}$. How good is this step as an approximation for the minimizer of $\Phi(x, \bar{\mu})$?

2. Consider the trust-region subproblem

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \;\; x^T g + \tfrac{1}{2}x^T B x \;\; \text{subject to} \;\; \|x\|_2 \leq \Delta$$

   (a) By observing that the constraints may be written as $x^T x \leq \Delta^2$, write down the logarithmic barrier function and its gradient and Hessian matrix.

   (b) Write down the first-order optimality conditions. How do they relate to Theorem 3.9?

# Exercises for Part 5.

1. Suppose that we wish to solve the equality-constrained quadratic program

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \;\; g^T x + \tfrac{1}{2} x^T B x \;\; \text{subject to} \;\; Ax = b,$$

where $g = -(1,1,1)^T$, $A = (1 \;\; 1 \;\; 0)$ and $b = 2$. Solve the problem when

(a)

$$B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

(b)

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \;\; \text{and}$$

(c)

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

2. Suppose that $B_k$ is positive definite, that $y_k$ are existing Lagrange multiplier estimates, and that $(s_k, y_{k+1})$ are the SQP search direction and its associated Lagrange multiplier estimates for the problem

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \;\; f(x) \;\; \text{subject to} \;\; c(x) = 0$$

at $x_k$. Then if $x_k$ is not a first-order critical point, show that $s_k$ is a descent direction for the *augmented Lagrangian* function

$$\Phi(x, y_k, \mu_k) = f(x) - y_k^T c(x) + \frac{1}{2\mu_k} \|c(x)\|_2^2$$

whenever

$$\mu_k \leq \frac{\|c(x_k)\|_2}{\|y_{k+1} - y_k\|_2}.$$

[This suggests that if we can adjust $y_k$ as the iteration proceeds so that $y_{k+1} - y_k \rightarrow 0$, we may not need $\mu_k$ to converge to zero. This is indeed the case for algorithms based upon the augmented Lagrangian function.]