Whenever calculations are needed to solve a problem, those calculations must be submitted as part of the homework assignment.

**Exercise 2.1.** Suppose that four foods are available to you: milk, chocolate chip cookies, muesli cereal, and pizza. Your daily diet right now consists of 2 pints of milk, 30 cookies, \(\frac{1}{2}\) bowl of muesli (which you dislike), and 5 slices of pizza. You have recently been informed by an authoritative source that you must consume more fiber and less fat, without adding any new foods to your diet.

You therefore wish to determine nonnegative quantities of the four foods available to you to make up a diet that maximizes your personal enjoyment while satisfying these conditions: (1) your daily fat intake must be less than or equal to 3,000; (2) you must consume at least 65 units of vitamin Z and 25 units of vitamin Y each day; and (3) you must consume at least 700 units of fiber each day.

The following table shows the amounts of vitamins Z and Y, fiber, fat, and enjoyment per indicated quantity of the four foods.

<table>
<thead>
<tr>
<th>Food</th>
<th>Z</th>
<th>Y</th>
<th>fiber</th>
<th>fat</th>
<th>enjoyment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milk (pint)</td>
<td>55</td>
<td>12</td>
<td>7</td>
<td>78</td>
<td>175</td>
</tr>
<tr>
<td>Cookies (each)</td>
<td>2</td>
<td>14</td>
<td>12</td>
<td>240</td>
<td>225</td>
</tr>
<tr>
<td>Muesli (bowl)</td>
<td>15</td>
<td>32</td>
<td>210</td>
<td>60</td>
<td>-200</td>
</tr>
<tr>
<td>Pizza (slice)</td>
<td>34</td>
<td>45</td>
<td>7</td>
<td>800</td>
<td>450</td>
</tr>
</tbody>
</table>

(a) Formulate a linear program of the form “minimize \(c^T x\) subject to \(Ax \geq b\)” whose solution will give you the most enjoyable diet satisfying conditions (1), (2), and (3), i.e., give \(A\), \(b\), and \(c\).

(b) Is your current diet feasible? Explain.

(c) Find a non-optimal vertex \(x_0\), explain how you found it, and compute \(c^T x_0\). Explain how you know that \(x_0\) is a vertex, and how you know that \(x_0\) is not optimal.

(d) Using your code that implements the all-inequality simplex method and starting with \(x_0\) from part (c), find the optimal solution \(x^*\). (See pages 56–57 of Notes 2.) Explain how you know that \(x^*\) is optimal and comment on any interesting features of \(x^*\). At the \(k\)th iteration of the simplex method, please give \(x_k\), \(\lambda_k\), \(p_k\), \(\alpha_k\), and \(c^T x_k\).

(e) Next, use your simplex code again, starting with \(x_0\) from part (c), to find the least enjoyable diet that satisfies the given constraints. Explain how the problem formulation differs from that in part (a). Confirm that your solution is optimal for this (reverse)
Exercise 2.2. Formulate the diet problem LP from Exercise 2.1 in standard form, namely give \( \bar{c}, \bar{A}, \) and \( \bar{b} \) such that the problem is to minimize \( \bar{c}^T \bar{x} \) subject to \( \bar{A} \bar{x} = \bar{b} \) and \( \bar{x} \geq 0 \).

Exercise 2.3. Consider the linear inequalities \( Ax \geq b \) with
\[
A = \begin{pmatrix}
4 & -1 \\
-1 & 0 \\
0 & 1
\end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix}
3 \\
-1 \\
2
\end{pmatrix}, \quad \text{(2.1)}
\]
for which there are no feasible points.

(a) Consider the following phase-1 linear program: minimize \( \hat{c}^T \hat{x} \) subject to \( \hat{A} \hat{x} \geq \hat{b} \), where
\[
\hat{A} = \begin{pmatrix}
4 & -1 & 1 \\
-1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix} \quad \hat{b} = \begin{pmatrix}
3 \\
-1 \\
2 \\
0
\end{pmatrix}, \quad \text{and} \quad \hat{c} = \begin{pmatrix}
0 \\
0 \\
1
\end{pmatrix}. \quad \text{(2.2)}
\]
State an interpretation of how the LP with \( \hat{A}, \hat{b}, \) and \( \hat{c} \) given by (2.2) is related to minimizing a measure of infeasibility for the constraints of (2.1).

(b) Find the optimal solution \( \hat{x} \) of the phase-1 LP specified by \( \hat{A}, \hat{b}, \) and \( \hat{c} \) of (2.2). What does \( \hat{x} \) tell us about the original constraints?

(c) Given a generic set of inequality constraints \( Ax \geq b \), where \( A \) is \( m \times n \), consider the phase-1 LP of minimizing \( \hat{c}^T \hat{x} \) subject to \( \hat{A} \hat{x} \geq \hat{b} \), with
\[
\hat{A} = \begin{pmatrix}
A & I_m \\
0_m & I_m
\end{pmatrix}, \quad \hat{b} = \begin{pmatrix}
b \\
0_m
\end{pmatrix}, \quad \text{and} \quad \hat{c} = \begin{pmatrix}
0_n \\
e
\end{pmatrix}. \quad \text{(2.3)}
\]
where \( \hat{A} \) is \( 2m \times (m+n) \), \( I_m \) is the \( m \)-dimensional identity matrix, \( 0_m \times n \) means an \( m \times n \) block of zeros, \( 0_m \) means a column of \( m \) zeros, and \( e \) is an \( m \)-vector whose components are all equal to 1, \( e = (1,1,\ldots,1)^T \). How is the solution of this LP related to the original constraints?

(d) Formulate and solve the specific phase-1 LP of minimizing \( \hat{c}^T \hat{x} \) subject to \( \hat{A} \hat{x} \geq \hat{b} \), where \( \hat{A}, \hat{b}, \) and \( \hat{c} \) are given by (2.3) and the entries of \( A \) and \( b \) are given in (2.1). Explain how to find a vertex \( \hat{x} \) where the first 2 components of \( \hat{x} \) are \( (1,2)^T \). Show the simplex steps that produce the optimal solution \( \hat{x}^* \) and describe its significance with respect to the original constraints.

Exercise 2.4. Consider the all-inequality linear program:
\[
\text{minimize } x_1 - \frac{1}{3}x_2 \text{ subject to } \begin{align*}
2 & \leq x_2, \\
1 & \leq 2x_2, \\
1 & \leq x_2, \\
1 & \leq x_2,
\end{align*} \quad \text{(2.4)}
\]
The constraints are shown in Figure 1 along with dotted lines showing two contours of the objective function. This same problem appears in Figure 3 in Notes 2 (page 57). The solution is $x^* = (-\frac{2}{3}, \frac{1}{3})^T$.

A simple primal-dual interior-point algorithm that uses Newton’s method with a backtracking line search is discussed in Notes 3 (pages 119–121). Apply your version of that algorithm to approximate the solution of the LP (2.4), starting with $\mu = 1$ and ending with $\mu = 10^{-4}$. Be sure to give the values of $\epsilon_1$, $\epsilon_2$, $\sigma$, and $\eta_s$ used in your numerical results. At the beginning of each outer iteration, print the value of $\|F(x, \lambda)\|$. At each inner iteration, print the current value of $\mu$, the iteration number, and the values of $x$ and $\lambda$, $\alpha_x$, $\alpha_\lambda$, $\alpha$, and $\|F_\mu(x, \lambda)\|$.

Please submit numerical results for two different starting points:

**Case 1.** $x_0 = \left(\frac{1}{2}, \frac{1}{2}\right)^T$ and $\lambda_0 = \left(\frac{3}{4}, \frac{1}{2}, \frac{1}{2}, \frac{1}{4}\right)^T$;

**Case 2.** Your own “interesting” choice of $x_0$ and $\lambda_0$. Please include an explanation of why your choice is interesting.

**Exercise 2.5.** This problem involves the famous Klee-Minty LP\(^1\) that produces worst-case behavior by the simplex method. The Klee-Minty LP can be described in numerous rescaled and rearranged forms, two of which appear in this problem.

(a) In one 2-d version of Klee-Minty, the problem is to maximize $10x_1 + x_2$ subject to $0 \leq x_1 \leq 1$, $x_2 \geq 0$, and $20x_1 + x_2 \leq 100$. Starting with the vertex $x_0 = (0, 0)^T$, execute the steps of the simplex method using the textbook rule. Show your calculations. What is the optimal vertex? How many steps were needed to reach it?

(b) Now consider a formulation of Klee–Minty in which $\epsilon$ is given such that $0 < \epsilon < \frac{1}{2}$. In $n$ dimensions, the constraints consist of the following pairs of upper and lower bounds,

$$\epsilon \leq x_1 \leq 1, \quad \epsilon x_{j-1} \leq x_j \leq 1 - \epsilon x_{j-1}, \quad j = 2, \ldots, n,$$

so that each successive variable is bounded above and below in terms of the preceding variable.

---

(i) Show that there are $2^n$ nondegenerate vertices for these constraints.

(ii) Suppose that we wish to maximize $x_n$ (the $n$-th variable) subject to these constraints. Show that the “highest” vertex is optimal, namely

\[
\begin{align*}
    x_1 &= \epsilon \\
    x_2 &= \epsilon x_1 = \epsilon^2 \\
    &\vdots \\
    x_{n-1} &= \epsilon x_{n-2} = \epsilon^{n-1} \\
    x_n &= 1 - \epsilon x_{n-1} = 1 - \epsilon^n.
\end{align*}
\]

(ii) Consider the vertex nearest to the origin, with $x_1 = \epsilon$, $x_2 = \epsilon^2$, \ldots, and $x_n = \epsilon^n$. Show that if the simplex method with the textbook deletion rule is started at this vertex, it reaches the optimal point in a single iteration.

Exercise 2.6. The 1-norm of a vector is the sum of the absolute values of its components. Given a nonzero $m \times n$ matrix $A$ and a nonzero $m$-vector $b$, suppose that we wish to find an $n$-vector $x$ that minimizes the 1-norm of the residual, so that our problem is

\[
\text{minimize} \quad \|Ax - b\|_1. \tag{2.5}
\]

(a) Express problem (2.5) as an all-inequality linear program of minimizing $\tilde{c}^T \tilde{x}$ subject to $A\tilde{x} \geq \tilde{b}$, i.e., define the variables $\tilde{x}$, the vector $\tilde{c}$, and the matrix and right-hand side of the constraints $A\tilde{x} \geq \tilde{b}$.

(b) Explain how your LP is connected to the original problem.

(c) Is there always a feasible point for the constraints of your LP? Explain.

Exercise 2.7.

(a) Given a vertex $\bar{x}$ for the constraints of a standard-form linear program ($Ax = b$, $x \geq 0$) in which $A$ has linearly independent rows, show that the columns of $A$ corresponding to positive components of $\bar{x}$ must be linearly independent.

[This result shows that, with the simplex strategy of including the indices of all positive components of a vertex in the basic set $B$, we guarantee that the corresponding columns of $A$ are linearly independent.]

(b) Consider the following example, in which $n = 3$, $m = 2$,

\[
A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad \text{and} \quad \bar{x} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.
\]

Is $\bar{x}$ a vertex? Explain why or why not.
Exercise 2.8. The following LP was constructed by Harold Kuhn (the second “K” in the KKT conditions) to illustrate that use of the textbook rules can lead to cycling:

\[
\begin{align*}
\text{minimize} & \quad -2x_1 - 3x_2 + x_3 + 12x_4 \\
\text{subject to} & \quad x_1, \ x_2, \ x_3, \ x_4 \geq 0 \\
& \quad 2x_1 + 9x_2 - x_3 - 9x_4 \geq 0 \\
& \quad -\frac{1}{3}x_1 - x_2 + \frac{1}{3}x_3 + 2x_4 \geq 0.
\end{align*}
\]

(a) Run the simplex method on this problem using the textbook rules, starting with \(x_0\) as the origin (a degenerate vertex where all six constraints are active) and taking the initial working set as \(W_0 = \{1, 2, 3, 4\}\). How many iterations occur before the initial working set is repeated (possibly in reordered form)?

(b) Now apply the simplex method, with the same initial point and working set, using Bland’s least-index choice rules, and demonstrate that the iterates move away from the origin at a subsequent iteration.

(c) Does the Kuhn LP have a bounded solution? Explain.

Exercise 2.9. Let \(q(x)\) be the quadratic function \(q(x) = c^T x + \frac{1}{2}x^T H x\). For each case given below, please do the following, including an explanation as well as relevant numerical calculations:

(i) Determine whether or not a stationary point \(x^*\) exists, and explain why or why not.

(ii) If a stationary point \(x^*\) exists, compute \(x^*\). Is \(x^*\) a minimizer of \(q\)? Explain why or why not.

(iii) If \(x^*\) is a minimizer of \(q\), is \(x^*\) unique? Explain. In either case, give the value of \(q(x^*)\);

(iv) If \(x^*\) is a stationary point but not a minimizer of \(q\), try to find a nonzero direction \(p\) such that \(q(x^* + \alpha p) < q(x^*)\) for all \(\alpha > 0\). Include an explanation of the procedure you used for computing \(p\).

(v) If no stationary point exists, state whether \(q(x)\) is unbounded below and explain your reasoning.

Case (a).
\[
c = \begin{pmatrix}
1 \\
3 \\
-7 \\
1
\end{pmatrix}, \quad H = \begin{pmatrix}
4 & 4 & 4 & 3 \\
4 & 7 & 3 & 3 \\
4 & 3 & 5 & 3 \\
3 & 3 & 3 & 3
\end{pmatrix}.
\]

Case (b).
\[
c = \begin{pmatrix}
1 \\
1 \\
1 \\
1
\end{pmatrix}, \quad H = \begin{pmatrix}
6 & 4 & 2 & 0 \\
4 & 7 & 3 & 3 \\
2 & 3 & 5 & 3 \\
0 & 3 & 3 & 3
\end{pmatrix}.
\]
Case (c).

\[ c = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad H = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 3 & 2 \\ 3 & 2 & 4 \end{pmatrix}. \]

Case (d).

\[ c = \begin{pmatrix} 0 \\ 1 \\ 5 \\ 5 \end{pmatrix}, \quad H = \begin{pmatrix} 11 & 7 & 3 & -2 \\ 7 & 6 & 5 & 2 \\ 3 & 5 & 7 & 6 \\ -2 & 2 & 6 & 7 \end{pmatrix}. \]

Case (e).

\[ c = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 3 \end{pmatrix}, \quad H = \begin{pmatrix} 3 & 3 & 0 & 3 \\ 3 & 7 & 4 & 3 \\ 0 & 4 & 6 & 2 \\ 3 & 3 & 2 & 5 \end{pmatrix}. \]

Exercise 2.10. This problem is intended to show that a pure Newton method for nonlinear equations can display interesting behavior even in two dimensions. We wish to find \( x^* \in \mathbb{R}^2 \) such that \( F(x^*) = 0 \), where

\[ F(x) = \begin{pmatrix} x_1^2 + 4x_2 \\ -\frac{1}{2}x_1 + 8x_2^2 \end{pmatrix}. \quad (2.6) \]

Write code that implements two methods: Method 1 is a pure Newton method, and Method 2 is Newton’s method with a simple backtracking line search using the merit function \( M(x) = \|F(x)\| \) and line search parameters \( \eta_s = .001 \) and \( \gamma_c = \frac{1}{2} \). (See pages 151 and 156 in Notes 4.) Your code should terminate after a specified maximum number of iterations \( \text{maxit} \) or when \( \|M_k\| \) is less than \( \text{ftol} \). At the \( k \)th iteration, print \( k, x_k, F_k, \) and \( \|F_k\| \), using scientific notation for the latter values and showing at least 6 significant digits. In the line search version, also print each trial value of the scalar \( \alpha_k \) (during the line search) and the value of the merit function.

(a) Find (by inspection or other simple means) two distinct zeros of \( F(x) \) as given in (2.6).
(b) Run Method 1 starting at \( x_0 = (1, -1)^T \) with \( \text{maxit} = 12 \). Are the iterates converging? To what point? What is the apparent rate of convergence of the merit function to zero?
(c) Starting at \( x_0 = (1.99, 0)^T \) with \( \text{maxit} = 9 \), run Method 1. Describe what happens to the iterates and to the merit function. Can you explain this behavior using properties of \( F \) and \( J \)?
(d) Starting at \( x_0 = (1.99, 0)^T \) with \( \text{maxit} = 15 \), run Method 2. What happens to the iterates compared with case (c)? How many non-unit values of \( \alpha_k \) occurred?
(e) Starting at \( x_0 = (1, 0)^T \) and with \( \text{maxit} = 8 \), run Method 1. Please describe and explain the behavior of the iterates.
(f) Starting at \( x_0 = (1, 0)^T \) and with \( \text{maxit} = 8 \), run Method 2. Comment on the chosen values of \( \{\alpha_k\} \) and the differences from the results of (e).