Chapter 1 Stochastic Linear Programming
1.1 Optimal Land Assignment - The Farmer’s Problem

A farmer has a total of 500 acres of land available for growing wheat, corn, and sugar beets. We denote by \( x_1, x_2, x_3 \) the amount of acres devoted to wheat, corn, and sugar beets. The planting costs per acre are 150, 230, and 260 US-Dollars for wheat, corn, and sugar beets. The farmer needs at least 200 tons \((T)\) of wheat and 240 T of corn for cattle feed which can be grown on the farm or bought from a wholesaler. We refer to \( y_1, y_2 \) as the amount of wheat and corn (in tons) purchased from the wholesaler. The purchase prices of wheat and corn per ton are 238 US-Dollars for wheat and 210 US-Dollars for corn. The amount of wheat and corn produced in excess will be sold at prices of 170 US-Dollars per ton for wheat and 150 US-Dollars per ton for corn. For sugar beets there is a quota on production which is 6000 T for the farmer. Any amount of sugar beets up to the quota can be sold at 36 US-Dollars per ton, the amount in excess of the quota is limited to 10 US-Dollars per ton. We denote by \( w_1 \) and \( w_2 \) the amount in tons of wheat and corn sold and by \( w_3, w_4 \) the amount in tons of sugar beets sold at the favorable price and the reduced price, respectively. The yield on the farmer’s land is a random variable \( \xi = (\xi_1, \xi_2, \xi_3)^T \) which can take on the realizations 3.0 T, 3.6 T, 24.0 T (above average), 2.5 T, 3.0 T, 20.0 T (average), 2.0 T, 2.4 T, 16.0 T (below average) per acre for wheat, corn, and sugar beets, each with probability \( 1/3 \). The farmer wants to maximize his profits.
<table>
<thead>
<tr>
<th></th>
<th>Wheat</th>
<th>Corn</th>
<th>Sugar Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yield (T/acre)</td>
<td>2.5</td>
<td>3.0</td>
<td>20.0</td>
</tr>
<tr>
<td>Planting cost ($/acre)</td>
<td>150</td>
<td>230</td>
<td>260</td>
</tr>
<tr>
<td>Purchase price ($/T)</td>
<td>238</td>
<td>210</td>
<td>–</td>
</tr>
<tr>
<td>Selling price ($/T)</td>
<td>170</td>
<td>150</td>
<td>36 (under 6000 T)</td>
</tr>
<tr>
<td>Minimum requirement (T)</td>
<td>200</td>
<td>240</td>
<td>–</td>
</tr>
</tbody>
</table>

Variables:

\[
\begin{align*}
\{x_1\} & \quad \text{amount of acres devoted to} \\
\{x_2\} & \quad \text{wheat} \\
\{x_3\} & \quad \text{corn} \\
\{y_1\} & \quad \text{sugar beets}
\end{align*}
\]

\[
\begin{align*}
\{y_2\} & \quad \text{amount of} \\
\\text{wheat} & \quad \text{purchased from a wholesaler} \\
\\text{corn} & \quad \text{favorable price} \\
\text{sugar beets} & \quad \text{sold at} \\
\text{sugar beets} & \quad \text{reduced price}
\end{align*}
\]

Total avail. land: 500 acres
Maximization of profits = Minimization of expenses

Linear program:

minimize \( \begin{align*} & 150x_1 + 230x_2 + 260x_3 + 238y_1 + 210y_2 - 170w_1 - 150w_2 - 36w_3 - 10w_4 \\ & \text{planting costs} \quad \text{purchase costs} \quad \text{income due to sales} \end{align*} \)

subject to:

\( \begin{align*} x_1 + x_2 + x_3 & \leq 500 \quad \text{(limitation of land)} \\ 2.5x_1 + y_1 - w_1 & \geq 200 \quad \text{(minimum requirement wheat)} \\ 3.0x_2 + y_2 - w_2 & \geq 240 \quad \text{(minimum requirement corn)} \\ w_3 + w_4 & \leq 20x_3 \quad \text{(max. yield sugar beets)} \\ w_3 & \leq 6000 \quad \text{(quota on sugar beets)} \end{align*} \)

\( x_1, x_2, x_3, y_1, y_2, w_1, w_2, w_3, w_4 \geq 0 \)
Solution of the linear program (average yield):

<table>
<thead>
<tr>
<th>Culture</th>
<th>Wheat</th>
<th>Corn</th>
<th>Sugar Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface (acres)</td>
<td>120</td>
<td>80</td>
<td>300</td>
</tr>
<tr>
<td>Yield (T)</td>
<td>300</td>
<td>240</td>
<td>6000</td>
</tr>
<tr>
<td>Purchases (T)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Sales (T)</td>
<td>100</td>
<td>–</td>
<td>6000</td>
</tr>
<tr>
<td>Maximum profit:</td>
<td>$ 118,600</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Solution of the linear program (above average yield):

<table>
<thead>
<tr>
<th>Culture</th>
<th>Wheat</th>
<th>Corn</th>
<th>Sugar Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface (acres)</td>
<td>183.33</td>
<td>66.67</td>
<td>250</td>
</tr>
<tr>
<td>Yield (T)</td>
<td>550</td>
<td>240</td>
<td>6000</td>
</tr>
<tr>
<td>Purchases (T)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Sales (T)</td>
<td>350</td>
<td>–</td>
<td>6000</td>
</tr>
<tr>
<td>Maximum profit:</td>
<td>$ 167,667</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Solution of the linear program (below average yield):

<table>
<thead>
<tr>
<th>Culture</th>
<th>Wheat</th>
<th>Corn</th>
<th>Sugar Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface (acres)</td>
<td>100</td>
<td>25</td>
<td>375</td>
</tr>
<tr>
<td>Yield (T)</td>
<td>200</td>
<td>60</td>
<td>6000</td>
</tr>
<tr>
<td>Purchases (T)</td>
<td>–</td>
<td>180</td>
<td>–</td>
</tr>
<tr>
<td>Sales (T)</td>
<td>–</td>
<td>–</td>
<td>6000</td>
</tr>
<tr>
<td>Maximum profit:</td>
<td>$ 59,950</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Stochastic Decision Problem (SDP)

Problem: (i) The farmer has to decide on the land assignment, i.e., determine $x_1, x_2, x_3$ (first stage decision variables), without knowing which scenario (average, above or below average) is going to happen.

(ii) The purchases $y_1, y_2$ and the sales $w_i; 1 \leq i \leq 4$, depend on the yield.

Conclusion: The decisions depend on the scenarios $1 \leq j \leq 3$:

- $j = 1$ (above average),
- $j = 2$ (average),
- $j = 3$ (below average)

New variables: $y_{ij}$ amount of \{wheat, corn\} purchased at scenario $j$

$w_{ij}$ amount of \{wheat, corn, sugar beets\} sold at scenario $j$

The variables $y_{ij}$ and $w_{ij}$ are called the second stage decision variables.
Extensive form of the stochastic decision problem

minimize $150x_1 + 230x_2 + 260x_3$ planting costs
$+ \frac{1}{3}(238y_{11} + 210y_{21})$ purchase costs at scenario $j = 1$
$+ \frac{1}{3}(238y_{12} + 210y_{22})$ purchase costs at scenario $j = 2$
$+ \frac{1}{3}(238y_{13} + 210y_{23})$ purchase costs at scenario $j = 3$

$- \frac{1}{3}(170w_{11} + 150w_{21} + 36w_{31} + 10w_{41})$ income due to sales $j = 1$
$- \frac{1}{3}(170w_{12} + 150w_{22} + 36w_{32} + 10w_{42})$ income due to sales $j = 2$
$- \frac{1}{3}(170w_{13} + 150w_{23} + 36w_{33} + 10w_{43})$ income due to sales $j = 3$
subject to:

\[
x_1 + x_2 + x_3 \leq 500 \quad \text{limitation of land}
\]

\[
3.0x_1 + y_{11} - w_{11} \geq 200 \quad \text{min. requirement wheat at scenario } j = 1
\]

\[
3.6x_2 + y_{21} - w_{21} \geq 240 \quad \text{min. requirement corn at scenario } j = 1
\]

\[
w_{31} + w_{41} \leq 24x_3 \quad \text{max. yield sugar beets at scenario } j = 1
\]

\[
w_{31} \leq 6000 \quad \text{quota on sugar beets at scenario } j = 1
\]

\[
2.5x_1 + y_{12} - w_{12} \geq 200 \quad \text{min. requirement wheat at scenario } j = 2
\]

\[
3.0x_2 + y_{22} - w_{22} \geq 240 \quad \text{min. requirement corn at scenario } j = 2
\]

\[
w_{32} + w_{42} \leq 20x_3 \quad \text{max. yield sugar beets at scenario } j = 2
\]

\[
w_{32} \leq 6000 \quad \text{quota on sugar beets at scenario } j = 2
\]

\[
2.0x_1 + y_{13} - w_{13} \geq 200 \quad \text{min. requirement wheat at scenario } j = 3
\]

\[
2.4x_2 + y_{23} - w_{23} \geq 240 \quad \text{min. requirement corn at scenario } j = 3
\]

\[
w_{33} + w_{43} \leq 16x_3 \quad \text{max. yield sugar beets at scenario } j = 3
\]

\[
w_{33} \leq 6000 \quad \text{quota on sugar beets at scenario } j = 3
\]
### Solution of the extensive form of the SDP

<table>
<thead>
<tr>
<th>First stage</th>
<th>Surface (acres)</th>
<th>Wheat</th>
<th>Corn</th>
<th>Sugar Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>above average</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>above average</td>
<td>Yield (T)</td>
<td>510</td>
<td>288</td>
<td>6000</td>
</tr>
<tr>
<td>above average</td>
<td>Purchases (T)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>above average</td>
<td>Sales (T)</td>
<td>310</td>
<td>48</td>
<td>6000</td>
</tr>
<tr>
<td>average</td>
<td>Yield (T)</td>
<td>425</td>
<td>240</td>
<td>5000</td>
</tr>
<tr>
<td>average</td>
<td>Purchases (T)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>average</td>
<td>Sales (T)</td>
<td>225</td>
<td>–</td>
<td>5000</td>
</tr>
<tr>
<td>below average</td>
<td>Yield (T)</td>
<td>340</td>
<td>192</td>
<td>4000</td>
</tr>
<tr>
<td>below average</td>
<td>Purchases (T)</td>
<td>–</td>
<td>48</td>
<td>–</td>
</tr>
<tr>
<td>below average</td>
<td>Sales (T)</td>
<td>140</td>
<td>–</td>
<td>4000</td>
</tr>
</tbody>
</table>

Expected maximum profit: $ 108,390

The difference ‘mean value - expected value’ of maximum profit ($ 7,016) is called the ‘expected value of perfect information’.
Reduced stochastic decision problem

**Feature:**
(i) The first stage decision (allocation of land) is made on the basis of average yield.
(ii) Purchases, sales, and yield still depend on the three scenarios.

**Homework:** Formulate the extensive form of the reduced SDP and solve it.

**Solution:** Expected max. profit for the reduced SDP: $107,240

The difference ‘expected max. profit SDP - expected max. profit reduced SDP’ ($1,150) is called the ‘value of the stochastic solution’.
1.2 The news vendor problem

Foundations of probability theory

A triple \((\Omega, \mathcal{F}, P)\) is called a probability space, if:

- \(\Omega\) is a set,
- \(\mathcal{F} \subseteq 2^{\Omega}\) is a \(\sigma\)-algebra, and
- \(P : \mathcal{F} \rightarrow [0, 1]\) is a probability measure.

\(\mathcal{F} \subseteq 2^{\Omega}\) is called a \(\sigma\)-algebra, if it is closed with respect to the complement, union, and intersection of countably many sets.

\(P : \mathcal{F} \rightarrow [0, 1]\) is called a probability measure, if:

(i) \(P(\emptyset) = 0,\ P(\Omega) = 1\),
(ii) \(P(\bigcup_{i \in I} E_i) = \bigcup_{i \in I} P(E_i),\ E_i \cap E_j = \emptyset,\ i \neq j \in I\).

A function \(X : \Omega \rightarrow \mathbb{R}\) is called a random variable, if:

\(\{\omega \mid X(\omega) \leq r\} \in \mathcal{F}\) for all \(r \in \mathbb{R}\).

Suppose a random variable \(X\) can take values \(x_k\) with probabilities \(p_k, 1 \leq k \leq n,\) such that \(\sum_{i=1}^{n} p_k = 1\). Then

\[
E(X) = \sum_{i=1}^{n} p_k x_k
\]

is called the expectation of \(X\).
Foundations of probability theory

A function \( f \) with \( \int_{-\infty}^{+\infty} f(x) \, dx = 1 \) is called a density function, if:

\[
P(X \leq x) = \int_{-\infty}^{x} f(s) \, ds \implies E(X) = \int_{-\infty}^{+\infty} xf(x) \, dx.
\]

A random variable \( X : \Omega \to \mathbb{R} \) with density function

\[
f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)
\]

is called \( N(\mu,\sigma^2) \)-normally distributed.

The special case \( \mu = 0, \sigma = 1 \) is called the standard normal distribution.
Tools from nonlinear optimization

Let $J : \mathbb{R}^n \rightarrow \mathbb{R}$ be a continuously differentiable convex functional and let $K \subset \mathbb{R}^n$ be a closed convex set. Consider the constrained minimization problem

$$(*) \quad J(u) = \inf_{v \in K} J(v).$$

If $u^* \in K$ solves $(*)$, then the following variational inequality holds true

$$\nabla J(u^*) \cdot (v - u^*) \geq 0, \quad v \in K.$$

Proof. If $v \in K$, then $u^* + \lambda(v - u^*) = \lambda v + (1 - \lambda)u^* \in K$, $\lambda \in [0, 1]$, since $K$ is convex. Moreover, it holds

$$0 \leq \frac{J(u^* + \lambda(v - u^*)) - J(u^*)}{\lambda} \xrightarrow{\lambda \rightarrow 0} \nabla J(u^*) \cdot (v - u^*).$$
Let $K$ be given by $K = \{ v \in \mathbb{R}^n \mid u_{\text{min}} \leq v \leq u_{\text{max}} \}$ for $u_{\text{min}} < u_{\text{max}}$.

Case I: $u^* = u_{\text{max}}$

Choose $v = u_{\text{max}} - w$ for $0 < w < u_{\text{max}} - u_{\text{min}}$, which implies $v \in K$. Then, it holds

$$-\nabla J(u^*)w \geq 0 \implies \nabla J(u^*)w \leq 0 \implies \nabla J(u^*) \leq 0.$$ 

If strict complementarity holds true, then $\nabla J(u^*) < 0$.

Case II: $u^* = u_{\text{min}}$

Choose $v = u_{\text{min}} + w$ for $0 < w < u_{\text{max}} - u_{\text{min}}$, which implies $v \in K$. Then, it holds

$$-\nabla J(u^*)w \geq 0 \implies \nabla J(u^*) \geq 0.$$ 

If strict complementarity holds true, then $\nabla J(u^*) > 0$.

Case III: $u_{\text{min}} < u^* < u_{\text{max}}$

Choose $v = u^* \pm w$ for $u^* - u_{\text{min}} < w < u_{\text{max}} - u^*$, which implies $v \in K$. Then, it holds

$$\pm \nabla J(u^*)w \geq 0 \implies \nabla J(u^*)w = 0 \implies \nabla J(u^*) = 0.$$
Stochastic linear program with continuous random variables: The news vendor problem

Publishing House → News Vendor → Customers

A news vendor buys \( x \leq x_{\text{max}} \) newspapers from a publishing house at a price of \( c \) per paper and sells the newspapers to customers at a price of \( q \) per paper. Unsold newspapers can be returned to the publisher at a price of \( r < c \).

The daily demand is described by a continuous random variable \( \xi \) with a known probability distribution \( F = F(\xi) \), i.e.,

\[
P(a \leq \xi \leq b) = \int_a^b dF(\xi), \quad \int_{-\infty}^{+\infty} dF(\xi) = 1.
\]

We assume a standard normal distribution, i.e.,

\[
dF(\xi) = f(\xi) \, d\xi \quad \text{with} \quad f(\xi) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\xi^2}{2}\right).
\]

The news vendor wants to maximize his profits.
The news vendor problem as a stochastic linear program

\[
\begin{align*}
\text{minimize} & \quad J(x) = cx + Q(x) \\
\text{subject to} & \quad 0 \leq x \leq x_{\text{max}},
\end{align*}
\]

where \( Q(x) = E_\xi Q(x, \xi) \) and \(-Q(x, \xi)\) is the expected profit on sales and returns.

Computation of \( Q(x, \xi) \) in case the demand \( \xi \) is known

We denote by \( y(\xi) \) the sold newspapers and by \( w(\xi) \) the remittents.

\[
\begin{align*}
\text{minimize} & \quad Q(x, \xi) = -qy(\xi) - rw(\xi) \\
\text{subject to} & \quad y(\xi) \leq \xi \quad \text{(sold newsp. less equal demand),} \\
& \quad y(\xi) + w(\xi) \leq x \quad \text{(sold newsp. + remittents less equal purchased newsp.),} \\
& \quad y(\xi), w(\xi) \geq 0.
\end{align*}
\]

The optimal solution is

\[
y^*(\xi) = \min(\xi, x), \quad w^*(\xi) = \max(x - \xi, 0)
\]

\[
\Rightarrow Q(x) = E_\xi (-q \min(\xi, x) - r \max(x - \xi, 0)).
\]
Computation of \( Q(x) = E_\xi Q(x, \xi) \) by means of the probability distribution \( F \)

\[
Q(x) = \int_{-\infty}^{+\infty} Q(x, \xi) \, dF(\xi) = \int_{-\infty}^{x} (-q\xi - r(x - \xi)) \, dF(\xi) + \int_{x}^{+\infty} (-qx) \, dF(\xi) = \\
-(q-r) \int_{-\infty}^{x} \xi \, dF(\xi) - rxF(x) - q\left( \int_{-\infty}^{+\infty} dF(\xi) - \int_{-\infty}^{x} dF(\xi) \right)
\]

Integration by parts:

\[
\int_{-\infty}^{x} \xi \, dF(\xi) = \xi F(\xi)|_{-\infty}^{x} - \int_{-\infty}^{x} F(\xi) \, d\xi = xF(x) - \int_{-\infty}^{x} F(\xi) \, d\xi
\]

It follows that

\[
Q(x) = -qx + (q-r) \int_{-\infty}^{x} F(\xi) \, d\xi
\]

\( Q(x) \) is differentiable with

\[
Q'(x) = -q + (q-r)F(x)
\]
From nonlinear optimization we know that the optimal solution $x^*$ satisfies the variational inequality

$$J'(x^*)(v - x^*) \geq 0, \quad v \in K := \{x \in \mathbb{R} | 0 \leq x \leq x_{\text{max}}\}$$

Strict complementarity conditions:

$$J'(x^*) < 0, \text{ if } x^* = x_{\text{max}}, \quad J'(x^*) > 0, \text{ if } x^* = 0, \quad J'(x^*) = 0, \text{ if } 0 < x^* < x_{\text{max}}$$

Computation of $J'(x)$

$$J'(x) = c + Q'(x) = c - q + (q - r)F(x) \implies$$

$$J'(x_{\text{max}}) < 0 \iff c - q + (q - r)F(x_{\text{max}}) < 0 \iff \frac{q - c}{q - r} > F(x_{\text{max}})$$

$$J'(0) > 0 \iff c - q + (q - r)F(0) > 0 \iff \frac{q - c}{q - r} < F(0)$$

Hence, we obtain

$$x^* = \begin{cases} 
0, & \text{if } (q - c)/(q - r) < F(0) \\
x_{\text{max}}, & \text{if } (q - c)/(q - r) > F(x_{\text{max}}) \\
F^{-1}((q - c)/(q - r)), & \text{otherwise}
\end{cases}$$
1.3 Two-stage stochastic program with recourse

\[ \begin{array}{ll}
\text{minimize} & c^T x + E_\xi Q(x, \xi) \\
\text{subject to} & Ax = b, \quad x \geq 0
\end{array} \]

where \( c \in \mathbb{R}^{n_1}, A \in \mathbb{R}^{m_1 \times n_1}, b \in \mathbb{R}^{m_1}. \) \( Q(x, \xi) \) is of the form

\[ Q(x, \xi) = \min \{ q^T y \mid Wy + Tx = h \} \]

where \( q \in \mathbb{R}^{n_2}, W \in \mathbb{R}^{m_2 \times n_2}, T \in \mathbb{R}^{m_2 \times n_1}, \) and \( h \in \mathbb{R}^{m_2}. \) \( q \) and \( h \) are random vectors.

The deterministic matrix \( W \) is called the \textit{recourse matrix} and the random matrix \( T \) is called the \textit{technology matrix}. \( Q(x) = E_\xi Q(x, \xi) \) is called the \textit{recourse function}.

The \textit{deterministic equivalent program} can be written as

\[ \begin{array}{ll}
\text{minimize} & c^T x + Q(x) \\
\text{subject to} & Ax = b, \quad x \geq 0
\end{array} \]
Feasible sets in case the random variable $\xi$ has $N \in \mathbb{N}$ realizations

$$K_1 := \{x \in \mathbb{R}^{n_1}_+ | Ax = b\}$$ is called the first stage feasible set.

$$K_2 := \{x \in \mathbb{R}^{n_1}_+ | Q(x) < \infty\}$$ is called the second stage feasible set.

$$K_2(\xi) := \{x \in \mathbb{R}^{n_1}_+ | Q(x, \xi) < \infty\}$$ is called the elementary second stage feasible set.

Let $\Sigma \subset \mathbb{R}^N$ be the support of $\xi$, i.e., $P(\{\xi \in \Sigma\} = 1$.

$$K_2^p := \{x \in \mathbb{R}^{n_1}_+ | \text{For all } \xi \in \Sigma \text{ there exists } y \geq 0 \text{ such that } Wy = h - Tx\}$$

Note that

$$K_2^p = \bigcap_{\xi \in \Sigma} K_2(\xi)$$

For the definition of the feasible sets in case of a continuous random variable $\xi$ we refer to the literature, since it requires advanced tools from stochastics.
Properties of second stage feasible sets

**Theorem.** For the second stage feasible sets it holds

(i) $K_2(\xi)$ is a closed convex polyhedron,

(ii) $K_2^p$ is a closed convex set,

(iii) If $\Sigma$ is finite, then $K_2^p = K_2$.

**Proof.** (i) is obvious. (ii) follows from the fact that the intersection of a finite number of closed convex sets is itself a closed convex set.

For the proof of (iii) we have

\[
x \in K_2 \implies Q(x) < \infty \implies Q(x, \xi) < \infty, \quad \xi \in \Sigma \implies x \in K_2(\xi), \quad \xi \in \Sigma \implies x \in K_2^p;
\]
\[
x \in K_2^p \implies Q(x, \xi) < \infty, \quad \xi \in \Sigma \implies Q(x) < \infty \implies x \in K_2.
\]

**Remark.** In case $\xi$ is a continuous random variable, see literature.
Example: Recourse function

\[
Q(x, \xi) := \begin{cases} 
\xi - x, & x \leq \xi \\
 x - \xi, & x \geq \xi 
\end{cases}
\]

Assume that \( \xi \) takes the values \( \xi_1 = 1, \xi_2 = 2, \xi_3 = 4 \), each with probability \( 1/3 \).

\[
Q(x) := \frac{1}{3} Q(x, \xi_1) + \frac{1}{3} Q(x, \xi_2) + \frac{1}{3} Q(x, \xi_3)
\]
The function \( Q(x) \) is not differentiable, but subdifferentiable.  
A convex function \( f : \mathbb{R}^n \to \mathbb{R} \) is said to be subdifferentiable at \( u \in \mathbb{R}^n \), if \( f \) has a continuous affine minorant \( \ell \) which is exact at \( u \). Hence, \( \ell \) has to be of the form
\[
\ell(v) = u^* \cdot (v - u) + f(u).
\]
The slope \( u^* \) of \( \ell \) is said to be the subgradient of \( f \) at \( u \) and the set \( \partial f(u) \) of all subgradients at \( u \) is called the subdifferential of \( f \) at \( u \). We have the following characterization:
\[
u^* \in \partial f(u) \iff u^* \cdot (v - u) + f(u) \leq f(v), \quad v \in \mathbb{R}^n.
\]

**Example:** The function \( f(x) = x \) has the subdifferential
\[
\partial f(x) = \begin{cases} 
-1, & x < 0 \\
[-1, +1], & x = 0 \\
+1, & x > 0
\end{cases}
\]

**Homework:** Determine the subdifferential of the recourse function \( Q(x) \) from the example.
Theorem (Properties of the second stage value function)
Assume that $\Sigma$ is finite and $Q(x, \xi) > -\infty$, $\xi \in \Sigma$. Then there holds

(i) $Q(x, \xi)$ is piecewise linear and convex in $(h, T)$.
(ii) $Q(x, \xi)$ is piecewise linear and concave in $q$.
(iii) $Q(x, \xi)$ is piecewise linear and concave in $x \in K_1 \cap K_2$.

Proof. For the proof of piecewise linearity we refer to the literature. In order to prove the convexity in $(h, T)$ we recall

$$Q(x, \xi) = \min \{ q^T y \mid Wy + Tx = h \}.$$

Hence, it suffices to show that

$$g(z) := \min \{ q^T y \mid Wy = z \}$$

is convex in $z$. 
Consider $z_1 \neq z_2$ and $z(\lambda) = \lambda z_1 + (1 - \lambda)z_2$, $\lambda \in [0, 1]$.

Let $y^*_1, y^*_2,$ and $y^*_\lambda$ be the optimal solutions with respect to $z_1, z_2, z(\lambda)$ and consider

$$y^*(\lambda) := \lambda y^*_1 + (1 - \lambda)y^*_2.$$ 

Obviously, $y^*(\lambda)$ is feasible and

$$g(z(\lambda)) = q^T y^*_\lambda \leq q^T y^*(\lambda) = \lambda q^T y^*_1 + (1 - \lambda)y^*_2 = \lambda g(z_1) + (1 - \lambda)g(z_2),$$

where the first inequality follows from the optimality of $y^*_\lambda$.

**Homework:** Prove concavity in $q$ and convexity in $x \in K_1 \cap K_2$.

**Remark.** For continuous random variables see the literature.
Optimality conditions for the Deterministic Equivalent Problem (DEP)

\[
\begin{align*}
\text{minimize} \quad & J(x) := c^T x + Q(x) \\
\text{subject to} \quad & Ax = b, \quad x \geq 0
\end{align*}
\]

**Theorem.** Assume that \( J(x) > -\infty \). Then, a feasible point \( x^* \in K_1 \) is an optimal solution of (DEP) if and only if there exist \( \lambda^* \in \mathbb{R}^{m_1}, \mu^* \in \mathbb{R}^{n_1}_+, (\mu^*)^T x^* = 0 \), such that

\[-c + A^T \lambda^* + \mu^* \in \partial Q(x^*),\]

where \( \partial Q(x^*) \) is the subdifferential of the recourse function at \( x^* \).

**Proof.** (DEP) is a constrained convex minimization problem with a closed convex constraint set. From Optimization Theory (Part I) we know that it is equivalent to the saddle point problem

\[
\inf_{x \in \mathbb{R}^{n_1}} \sup_{\lambda \in \mathbb{R}^{m_1}, \mu \in \mathbb{R}_+^{n_1}} L(x, \lambda, \mu),
\]

where \( L(x, \lambda, \mu) \) stands for the Lagrangian

\[
L(x, \lambda, \mu) = J(x) - \lambda^T (Ax - b) - \mu^T x.
\]

One of the optimality conditions for the saddle point problem is

\[
0 \in \partial_x L(x^*, \lambda^*, \mu^*) = \partial J(x^*) - A^T \lambda^* - \mu^* = c + \partial Q(x^*) - A^T \lambda^* - \mu^*.
\]
Homework. Consider the second stage decision problem

\[
\text{minimize} \quad E_{\xi} Q(x, \xi)
\]
\[
Q(x, \xi) := \begin{cases} 
\xi - x, & x \leq \xi \\
-x, & x \geq \xi 
\end{cases}
\]

where \(\xi\) takes the values \(\xi_1 = 1, \xi_2 = 2, \xi_3 = 4\), each with probability \(1/3\).

Compute the optimal solution \(x^*\) by using the previous theorem.
Representation of the subdifferential of the recourse function

For the recourse function in the previous homework it holds

\[ \partial Q(x) = \partial (E_\xi Q(x, \xi)) = E_\xi(\partial Q(x, \xi)). \]

However, the above relationship does not hold true in general. In general, it holds

\[ \partial Q(x) = E_\xi(\partial Q(x, \xi)) + \text{rec}(\partial Q(x)), \]

where \( \text{rec}(\partial Q(x)) \), stands for the recession cone

\[ \text{rec}(\partial Q(x)) = \{ v \mid \text{For all } u \in \partial Q(x), \lambda \geq 0: u + \lambda v \in \partial Q(x) \}. \]

Example. Consider

\[ Q(x) = \begin{cases} 0, & x \leq 0 \\ +\infty, & x > 0 \end{cases}. \]

\( Q(x) \) is the indicator function of the closed convex set \( K = \{ x \mid x \geq 0 \} \).

\[ \partial Q(x) = \begin{cases} 0, & x < 0 \\ [0, \infty], & x = 0 \end{cases}, \quad \text{rec}(\partial Q(0)) = [0, \infty]. \]
Properties of the recession cone

Lemma. If the convex set $K$ is bounded, then $\text{rec}(K) = \{0\}$. The converse is not true: Consider

$$K := \{(x, y) \mid 0 \leq x < 1, \ y \geq 1\} \cup \{(x, y) \mid 0 \leq x \leq 1, \ 0 \leq y \leq 1\}.$$ 

$K$ is not bounded, but $\text{rec}(K) = \{0\}$.

Examples.

(i) $K := \{(x, y) \mid |x| \leq y\} \implies \text{rec}(K) = K$

(ii) $K := \{(x, y) \mid |x|^n \leq y, \ n > 1\} \implies \text{rec}(K) = \{(0, y) \mid y \geq 0\}$
Decomposition of the subdifferential

**Theorem.** For \( x \in K_1 \cap K_2 \) it holds

\[
\partial Q(x) = E_\omega(\partial Q(x, \xi(\omega))) + N(K_2, x),
\]

where \( N(K_2, x) \) is the normal cone of \( K_2 \) at \( x \), i.e.,

\[
\partial N(K_2, x) = \{ v \in \mathbb{R}^{n_1} \mid v^T y \leq 0 \text{ for all } y \in \mathbb{R}^{n_1} \text{ such that } x + y \in K_2 \}.
\]

**Proof.** From the previous theorem we know

\[
\partial Q(x) = E_\omega \partial Q(x, \xi(\omega)) + \text{rec}(\partial Q(x)).
\]

We recall

\[
\text{u} \in \partial Q(x) \iff Q(x) \text{ is finite and } (+) \text{ u}^T (w - x) + Q(x) \leq Q(w) \text{ for all } w \in \mathbb{R}^{n_1}.
\]

a) Choose \( w = x + y \) in \((+)\) \( \implies \)

\[
(o) \quad u^T y + Q(x) \leq Q(x + y).
\]
b) \( u + \lambda v \in \partial Q(x) \iff \)

\[ (u + \lambda v)^T(w - x) + Q(x) \leq Q(w), \quad w \in \mathbb{R}^{n_1}. \]

Choose again \( w = x + y \Rightarrow \)

\[ (u + \lambda v)^Ty + Q(x) \leq Q(x + y) \iff \]

\[ \text{rec}(\partial Q(x)) = \{ v \in \mathbb{R}^{n_1} | y^T(u + \lambda v) + Q(x) \leq Q(x + y), \lambda \geq 0, y \in \mathbb{R}^{n_1} \}. \]

Now:

\[ y^Tu + \lambda y^Tv + Q(x) \leq Q(x + y) \]

\[ y^Tu + Q(x) \leq Q(x + y) \text{ due to (o)} \]

\[ \Rightarrow v \in \text{rec}(\partial Q(x)) \iff y^Tv \leq 0 \text{ for all } y \text{ such that } \frac{Q(x + y)}{x+y \in K_2} < \infty. \]
1.4 The L-shaped algorithm

Assumption: ξ has K possible realizations with probabilities \( p_k, 1 \leq k \leq K \).

Feature: The algorithm is an iterative method consisting of feasibility cuts and optimality cuts.

Motivation feasibility cuts: Test second stage feasibility of an iterate \( x'' \). Does \( x'' \in K_2 \) hold true?

\[
x'' \in K_2 \iff x'' \in \{ x \in \mathbb{R}^n_1 | \text{There exists } y \geq 0 \text{ such that } Wy = h_k - T_k x, 1 \leq k \leq K \}.
\]

We define the positive cone of \( W \) as follows:

\[
\text{pos}(W) := \{ t \in \mathbb{R}^{m_2} | t = Wy, y \geq 0 \}.
\]

Hence, we have

\[(*)_1 \quad h_k - T_k x'' \in \text{pos}(W) \text{ for all } 1 \leq k \leq K, \]
\[(*)_2 \quad \text{There exists } \bar{k} \in \{1, \ldots, K\} \text{ such that } h_{\bar{k}} - T_{\bar{k}} x'' \notin \text{pos}(W).\]

If \((*)_1\) is satisfied: \( x'' \) is second stage feasible.

If \((*)_2\) holds true: \( x'' \) is not second stage feasible.
Separation of convex cones

If $x^\nu$ is not second stage feasible, there exists a hyperplane separating $h \bar{\kappa} - T \bar{\kappa} x^\nu$ and $\text{pos}(W)$. In particular, there exists $\sigma \in \mathbb{R}^m$ such that

$$\sigma^T t \leq 0 \text{ for all } t \in \text{pos}(W) \text{ and } \sigma^T (h \bar{\kappa} - T \bar{\kappa} x^\nu) > 0.$$
How to construct that hyperplane?

We need the relationship between **primal and dual linear programs** known from Optimization I.

**Primal LP:** minimize \( c^T x \)
subject to \( Ax = b, \ x \geq 0 \)

**Dual LP:** maximize \( b^T \sigma \)
subject to \( A^T \sigma + s = c, \ s \geq 0 \)

where \( A \in \mathbb{R}^{m \times n}, c \in \mathbb{R}^n, b \in \mathbb{R}^m, \sigma \in \mathbb{R}^m, s \in \mathbb{R}^n \).

**Theorem.** If \( x^* \in \mathbb{R}^n \) and \( \sigma^* \in \mathbb{R}^m \) are the minimizer/maximizer, then it holds
\[
b^T \sigma^* = c^T x^*.
\]
Construction of the hyperplane

For \(1 \leq k \leq K\) consider the linear programs

**Primal (LP)\(_k\):**

\[
\begin{align*}
\text{minimize} & \quad \tilde{J}(y, v^+, v^-) := e^T v^+ + e^T v^- \quad (e := (1, \cdots, 1)^T) \\
\text{subject to} & \quad Wy + v^+ - v^- = h_k - T_k x^\nu, \quad y, v^+, v^- \geq 0
\end{align*}
\]

**Standard form of (LP)\(_k\):** Set

\[
\begin{align*}
z := (y, v^+, v^-)^T, \quad c := (0, e, e)^T, \quad A = (W \mid I \mid - I), \quad b := h_k - T_k x^\nu \quad \Rightarrow \\
\text{Primal (LP)\(_k\):} & \quad \text{minimize} \quad c^T z \quad \text{subject to} \quad Az = b, \quad z \geq 0 \\
\text{Dual (LP)\(_k\):} & \quad \text{maximize} \quad b^T \sigma \quad \text{subject to} \quad A^T \sigma + s = c, \quad s \geq 0
\end{align*}
\]

\[
A^T \sigma + s = c \quad \Leftrightarrow \quad \begin{pmatrix} W^T \\ I \\ -I \end{pmatrix} \sigma + \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{pmatrix} 0 \\ e \\ e \end{pmatrix}
\]

\[
\Rightarrow W^T \sigma + s_1 \geq 0 \quad \Rightarrow \quad (*) \quad W^T \sigma \leq 0
\]
Second stage feasibility

Let \( y_k' \in \mathbb{R}^{m_2} \) and \( \sigma_k' \in \mathbb{R}^{m_2} \), \( 1 \leq k \leq K \), be the solutions of Primal (LP)\(_k\) and Dual (LP)\(_k\), respectively. From the previous theorem we know

\[
\tilde{J}(y_k', v^+, v^-) = (\sigma_k')^T(h_k - T_kx''), \ 1 \leq k \leq K. 
\]

Case I: Second stage feasibility

\[
\tilde{J}(y_k', v^+, v^-) = 0 \text{ for all } 1 \leq k \leq K \quad \implies \\
(\sigma_k')^T(h_k - T_kx'') = 0, \ (\sigma_k')^TW \leq 0 \text{ for all } 1 \leq k \leq K \quad \implies \\
h_k - T_kx'' \in \text{pos}(W) \text{ for all } 1 \leq k \leq K \quad \implies \quad x'' \in K_2. 
\]

Case II: Second stage infeasibility

\[
\tilde{J}(y_k', v^+, v^-) > 0 \text{ for some } 1 \leq k \leq K \quad \implies \\
(\sigma_k')^T(h_k - T_kx'') > 0, \ (\sigma_k')^TW \leq 0 \text{ for some } 1 \leq k \leq K \quad \implies \\
h_k - T_kx'' \notin \text{pos}(W) \quad \implies \quad x'' \notin K_2. 
\]
Feasibility cuts

Since we want the subsequent iterates to be situated on the correct side of the hyperplane, i.e.,

\[(\sigma_{k}^{\nu})^T(h_k - T_k x) \leq 0,\]

we impose the feasibility cuts

\[\begin{aligned}
\underbrace{(\sigma_{k}^{\nu})^T T_k}_{D_k} x &\geq \underbrace{(\sigma_{k}^{\nu})^T h_k}_{d_k}.
\end{aligned}\]
Optimality cuts

Motivation optimality cuts: Test optimality of the iterate \( x' \)

First observation: The deterministic equivalent problem (DEP) can be written as:

\[
(\text{DEP}') \quad \begin{align*}
\text{minimize} & \quad c^T x + \Theta, \quad Q(x) \leq \Theta \\
\text{subject to} & \quad x \in K_1 \cap K_2
\end{align*}
\]

For \( 1 \leq k \leq K \) consider the linear programs

Primal \((LP)_k\) \quad \begin{align*}
\text{minimize} & \quad \hat{J}_k(y) = q^T y \\
\text{subject to} & \quad W y = h_k - T_k x' , \quad y \geq 0
\end{align*}

The associated dual programs read as follows

Dual \((LP)_k\) \quad \begin{align*}
\text{minimize} & \quad \pi_k^T (h_k - T_k x') \\
\text{subject to} & \quad W^T \pi_k + s = q, \quad \pi_k, s \geq 0
\end{align*}
Lemma. Representation of subgradients.

Let $\pi_k$ be the optimal solutions of the Dual $(LP)_k$. Then it holds

$$-(\pi_k)^T T_k \in \partial Q(x', \xi_k).$$

(*)

Proof. Homework.

Let $\pi_k$ be the optimal solutions of the Dual $(LP)_k$. The duality theory of linear programs implies

$$Q(x', \xi_k) = (\pi_k)^T (h_k - T_k x'), \quad 1 \leq k \leq K.$$ (+)

$$v \in \partial Q(x', \xi_k) \iff v^T (x - x') + Q(x', \xi_k) \leq Q(x, \xi_k), \quad x \in \mathbb{R}^n.$$  

Together with (*) from the previous lemma this implies

$$(o) \quad (\pi_k)^T T_k (x' - x) + Q(x', \xi_k) \leq Q(x, \xi_k), \quad x \in \mathbb{R}^n.$$  

Using (+) in (o) yields

$$Q(x, \xi_k) \geq (\pi_k)^T (h_k - T_k x).$$
It follows that
\[ Q(x^\nu) = E(\pi^\nu)^T(h - Tx^\nu) = \sum_{k=1}^{K} p_k(\pi_k^\nu)^T(h_k - T_k x^\nu), \]
\[ Q(x) = E(\pi^\nu)^T(h - Tx) = \sum_{k=1}^{K} p_k(\pi_k^\nu)^T(h_k - T_k x). \]

Hence, a pair \((x, \Theta)\) is feasible for \((\text{DEP})'\), if and only if
\[ \Theta \geq Q(x) \geq E(\pi^\nu)^T(h - Tx). \]

Moreover, a pair \((x^\nu, \Theta^\nu)\) is optimal for \((\text{DEP})'\), if and only if
\[ Q(x^\nu) = \Theta^\nu = E(\pi^\nu)^T(h - Tx^\nu). \]

Check for optimality (assuming that an iterate \(\Theta^\nu\) is available

Compute \( J^\nu := \sum_{k=1}^{K} p_k(\pi_k^\nu)^T(h_k - T_k x^\nu) = Q(x^\nu). \)

Case I: \( \Theta^\nu \geq J^\nu \) Stop the iteration: \( x^\nu \) is optimal.

Case II: \( \Theta^\nu < J^\nu \) \( x^\nu \) is not optimal.
Impose the optimality cut

\[ \sum_{k=1}^{K} p_k (\pi_k^\nu)^T T_k x + \Theta \geq \sum_{k=1}^{K} p_k (\pi_k^\nu)^T h_k \]

\[ \underbrace{\sum_{k=1}^{K} p_k (\pi_k^\nu)^T T_k x + \Theta}_{E_s \geq} \underbrace{\sum_{k=1}^{K} p_k (\pi_k^\nu)^T h_k}_{e_s} \]
Implementation of the L-shaped method

\( \nu := \) iteration counter  \( r : \) counter for feasibility cuts  \( s : \) counter for optimality cuts

Step 0: Set \( \nu = r = s = 0 \)

Step 1: Set \( \nu = \nu + 1 \) and solve the linear program

\[
\begin{align*}
\text{minimize} & \quad z = c^T x + \Theta \\
\text{subject to} & \quad A x = b \\
& \quad D_\ell x \geq d_\ell, \quad \ell = 1, \ldots, r \\
& \quad E_\ell x + \Theta \geq e_\ell, \quad \ell = 1, \ldots, s \\
& \quad x \geq 0, \quad \Theta \in \mathbb{R} \\
\end{align*}
\]

Let \( (x^\nu, \Theta^\nu) \) be the optimal solution (if \( s=0, \) set \( \Theta^\nu = -\infty \))
Step 2: Feasibility cuts

For $1 \leq k \leq K$ solve the linear programs

$$\begin{align*}
\text{minimize} & \quad \tilde{J}_k : = e^T v^+ + e^T v^- \\
\text{subject to} & \quad W y + I v^+ - I v^- = h_K - T_K x'' \\
& \quad y, v^+, v^- \geq 0
\end{align*}$$

Case I: If $\tilde{k} \in \{1, \cdots, K\}$ such that $\tilde{J}_{\tilde{k}} > 0$, let $\sigma_k^{\nu}$ be the optimal solution of the associated dual program and set

$$D_{r+1} := (\sigma_k^{\nu})^T T_k, \quad d_{r+1} := (\sigma_k^{\nu})^T h_{\tilde{k}}$$

Set $r = r + 1$ and go back to Step 1.

Case II: If $\tilde{J}_k^{\nu} = 0$, for all $1 \leq k \leq K$, go to Step 3.
Step 3: Optimality cuts

For $1 \leq k \leq K$ solve the linear programs

$$\begin{align*}
\text{minimize} & \quad q^T y \\
\text{subject to} & \quad Wy = h_k - T_k x'' \\
& \quad y \geq 0
\end{align*}$$

Let $\pi_k''$ be the optimal solutions of the associated dual programs and set

$$E_{s+1} := \sum_{k=1}^{K} p_k(\pi_k'')^T T_k, \quad e_{s+1} := \sum_{k=1}^{K} p_k(\pi_k'')^T h_k$$

Case I: $\Theta'' \geq e_{s+1} - E_{s+1} x''$

Stop the algorithm: $x''$ is an optimal solution.

Case II: $\Theta'' < e_{s+1} - E_{s+1} x''$

Set $s = s + 1$ and go back to Step 1.
Example: Feasibility cuts

Consider the following two-stage stochastic linear program

\[
\begin{align*}
\text{minimize} & \quad 3x_1 + 2x_2 + \mathbb{E}_\xi \min(-15y_1, 12y_2) \\
\text{subject to} & \quad 3y_1 + 2y_2 \leq x_1 \\
& \quad 2y_1 + 5y_2 \leq x_2 \\
& \quad 0.8\xi_1 \leq y_1 \leq \xi_1 \\
& \quad 0.8\xi_2 \leq y_2 \leq \xi_2 \\
& \quad x_1 \geq 0, \ y_i \geq 0, \ 1 \leq i \leq 2
\end{align*}
\]

It represents an investment decision problem in two resources \(x_1, x_2\), where in the second stage decision 80% of the demand has to be satisfied.

We assume \((\xi_1, \xi_2) = (4, 4)\) and \((\xi_1, \xi_2) = (6, 8)\), with probability \(1/2\) each.
By introducing slack variables $W_i \geq 0, 1 \leq i \leq 6$, we write the constraints in the form

$$
\begin{bmatrix}
3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\
2 & 5 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \\
\end{bmatrix}
\begin{bmatrix}
y_1 \\
y_2 \\
w_1 \\
w_2 \\
w_3 \\
w_4 \\
w_5 \\
w_6 \\
\end{bmatrix}
= 
\begin{bmatrix}
0 \\
0 \\
0 \\
-0.8\xi_1 \\
0 \\
0 \\
-0.8x_2 \\
0 \\
\end{bmatrix}
= h
\begin{bmatrix}
-1 & 1 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
= T
\begin{bmatrix}
x_1 \\
x_2 \\
\end{bmatrix}
$$

**Practical Homework 1.** Write a code that implements the L-shaped method. Use the code to perform feasibility cuts for the example above starting from $x_1 = x_2 = 0$. The delivery should contain the used code and a documentation of all steps of the L-shaped method.
Example: Optimality cuts

Consider the following two-stage stochastic linear program

\[
\begin{align*}
\text{minimize} & \quad Q(x) \\
\text{subject to} & \quad 0 \leq x \leq 10 \\
Q(x, \xi) & = \begin{cases} 
\xi - x , & x \leq \xi \\
x - \xi , & x \geq \xi 
\end{cases}
\end{align*}
\]

Let \( \xi_1 = 1, \xi_2 = 2, \xi_3 = 4, \) be possible realizations of \( \xi \) with probabilities \( p_k = 1/3, 1 \leq k \leq 3. \)

\[
Q(x, \xi_k) = \min \{ q_k^T y \mid W y = h_k - T_k x, \ y \geq 0 \}
\]

Hence, we have

\[
q_k = 1, \ W = 1, \ T_k = \begin{cases} 
1 , & x \leq \xi_k \\
-1 , & x > \xi_k 
\end{cases}, \ h_k = \begin{cases} \xi_k , & x \leq \xi_k \\
-\xi_k , & x > \xi_k \end{cases}
\]

Initialization: Set \( \nu = r = s = 0, \ \Phi_r = 0, \ d_r = 0, \ \Phi_s = 0, \ e_s = 0, \ x^1 = 0, \ \Theta^1 = -\infty \)
Iteration 1: $\nu = 1$

Step 2 (Feasibility cuts): Since $x^1 = 0$ is feasible, no feasibility cuts.

Step 3 (Optimality cuts): Consider the minimization problems

\[
\begin{align*}
\text{minimize} & \quad \tilde{J}^1_k(y) = q_k^Ty = y \\
\text{subject to} & \quad y = h_k - T_k x^1 = h_k, \ 1 \leq k \leq 3, \ y \geq 0 
\end{align*}
\]

The associated dual linear programs read

\[
\begin{align*}
\text{maximize} & \quad h_k^T\pi_k^1 \\
\text{subject to} & \quad \pi_k^1 + s = 1, \ s \geq 0 
\end{align*}
\]

The solution is $y^1 = (1, 2, 4)^T$ with $\pi_k^1 = 1, 1 \leq k \leq 3$. This implies

\[
E_1 = \sum_{k=1}^{3} p_k \pi_k^1 T_k = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1, \ e_1 = \sum_{k=1}^{3} p_k \pi_k^1 h_k = \frac{1}{3} + \frac{2}{3} + \frac{4}{3} = \frac{7}{3}, \ J^1 = e_1 - E_1 x^1 = e_1 = \frac{7}{3} \quad \Theta^1 = -\infty \nless J^1.
\]

Set $s = 1$ and continue.
Iteration 2: \( \nu = 2 \)

Step 1: Solve the minimization problem

\[
\begin{align*}
\text{minimize} \quad & \Theta \\
\text{subject to} \quad & \Theta \geq \frac{7}{3} - x, \ 0 \leq x \leq 10, \ \Theta \in \mathbb{R}
\end{align*}
\]

The solution is \( x^2 = 10, \Theta^2 = -\frac{23}{3} \).

Step 2 (Feasibility cuts): Since \( x^2 = 10 \) is feasible, no feasibility cuts.

Step 3 (Optimality cuts): Consider the minimization problems

\[
\begin{align*}
\text{minimize} \quad & \tilde{J}_k^2(y) = q_k^T y = y \\
\text{subject to} \quad & y = h_k - T_k \begin{pmatrix} x_2^2 \end{pmatrix} = -\xi_k + x^2, \ 1 \leq k \leq 3, \ y \geq 0
\end{align*}
\]

The associated dual linear programs read

\[
\begin{align*}
\text{maximize} \quad & (-\xi_k + x^2)^T \pi_k^2 \\
\text{subject to} \quad & \pi_k^2 + s = 1, \ s \geq 0
\end{align*}
\]

The solution is \( y^2 = (9, 8, 6)^T \) with \( \pi_k^2 = 1, 1 \leq k \leq 3 \). This implies

\[
E_2 = \sum_{k=1}^{3} p_k \pi_k^2 T_k = -\begin{pmatrix} \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \end{pmatrix} = -1, \ e_2 = \sum_{k=1}^{3} p_k \pi_k^2 h_k = -\begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & -\frac{4}{3} \end{pmatrix} = -\frac{7}{3},
\]

\[
J^2 = e_2 - E_2 x^2 = \frac{23}{3} \quad \Theta^2 = -\frac{23}{3} \not\geq J^2.
\]

Set \( s = 2 \) and continue.
Iteration 3: \( \nu = 3 \)

Step 1: Solve the minimization problem

\[
\begin{align*}
\text{minimize} & \quad \Theta \\
\text{subject to} & \quad \Theta \geq \frac{7}{3} - x, \quad \Theta \geq x - \frac{7}{3}, \quad 0 \leq x \leq 10, \quad \Theta \in \mathbb{R}
\end{align*}
\]

The solution is \( x^3 = \frac{7}{3}; \Theta^3 = 0 \).

Step 2 (Feasibility cuts): Since \( x^3 = \frac{7}{3} \) is feasible, no feasibility cuts.

Step 3 (Optimality cuts): Consider the minimization problems

\[
\begin{align*}
\text{minimize} & \quad \tilde{J}^3_k(y) = q_k^T y = y \\
\text{subject to} & \quad y = h_k - T_k x^3 \quad = (4/3, 1/3, 5/3)^T, \quad y \geq 0
\end{align*}
\]

The associated dual linear programs read

\[
\begin{align*}
\text{maximize} & \quad (h_k - T_k x^3)^T \pi_k^3 \\
\text{subject to} & \quad \pi_k^3 + s = 1, \quad s \geq 0
\end{align*}
\]

The solution is \( y^3 = (4/3, 1/3, 5/3)^T \) with \( \pi_k^3 = 1, 1 \leq k \leq 3 \). This implies

\[
E_3 = \sum_{k=1}^{3} p_k \pi_k^3 T_k = -\frac{1}{3}, \quad e_3 = \sum_{k=1}^{3} p_k \pi_k^3 h_k = \frac{1}{3},
\]

\[
J^3 = e_3 - E_3 x^3 = \frac{10}{9}, \quad \Theta^3 = 0 < J^3.
\]

Set \( s = 3 \) and continue.
**Homework.** Continue with the iteration until you find an optimal solution.