Chapter 1

Stochastic Linear and Nonlinear Programming

1.1 Optimal land usage under stochastic uncertainties

1.1.1 Extensive form of the stochastic decision program

We consider a farmer who 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 of land devoted to wheat, corn and sugar beets, respectively.

The planting costs per acre are \$ 150, \$ 230, and \$ 260 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 resp. corn (in tons) purchased from a wholesaler. The purchase prices of wheat resp. corn per ton are \$ 238 for wheat and \$ 210 for corn.

The amount of wheat and corn produced in excess will be sold at prices of \$ 170 per ton for wheat and \$ 150 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 per ton, the amount in excess of the quota is limited to \$ 10 per ton. We denote by w_1 and w_2 the amount in tons of wheat resp. corn sold and by w_3, w_4 the amount of sugar beets sold at the favorable price and the reduced price, respectively.

The farmer knows that the average yield on his land is 2.5 T, 3.0 T and 20.0 T per acre for wheat, corn and sugar beets.

The data are shown in Table 1.

TABLE 1. Data for optimal land usage

	Wheat	Corn	Sugar Beets
Yield (T/acre)	2.5	3.0	20.0
Planting cost (\$/acre)	150	230	260
Purchase price (\$/T)	238	210	–
Selling price (\$/T)	170	150	36 (under 6000 T)
			10 (above 6000 T)
Minimum requirement (T)	200	240	–
Total available land: 500 acres			

The farmer wants to maximize his profit. Based on the above data, this amounts to the solution of the linear program:

$$
\begin{aligned}
(1.1) \qquad & \text{minimize} && (150x_1 + 230x_2 + 260x_3 + 238y_1 - 170w_1 \\
& && + 210y_2 - 150w_2 - 36w_3 - 10w_4) \\
& \text{subject to} && x_1 + x_2 + x_3 \le 500 , \\
& && 2.5x_1 + y_1 - w_1 \ge 200 , \\
& && 3.0x_2 + y_2 - w_2 \ge 240 , \\
& && w_3 + w_4 \le 20x_3 , \\
& && w_3 \le 6000 , \\
& && x_1, x_2, x_3, y_1, y_2, w_1, w_2, w_3, w_4 \ge 0 .
\end{aligned}
$$

The solution of (1.1) is shown in Table 2.

TABLE 2. Solution of the linear program ('average yields')

Culture	Wheat	Corn	Sugar Beets
Surface (acres)	120	80	300
Yield (T)	300	240	6000
Purchases (T)	–	–	–
Sales (T)	100	–	6000
Maximum profit: $ 118,600			

The yield is sensitive to, e.g., wheather conditions. We refer to to the previously determined optimal solution as that one based on 'average yields' and consider two more scenarios, namely 'above average yields' and 'below average yields' by a margin of ± 20 %. The associated optimal solutions are depicted in Table 3 and Table 4.

TABLE 3. Solution of the linear program ('above average yields')

Culture	Wheat	Corn	Sugar Beets
Surface (acres)	183.33	66.67	250
Yield (T)	550	240	6000
Purchases (T)	–	–	–
Sales (T)	350	–	6000
Maximum profit: $ 167,667			

TABLE 4. Solution of the linear program ('below average yields')

Culture	Wheat	Corn	Sugar Beets
Surface (acres)	100	25	375
Yield (T)	200	60	6000
Purchases (T)	–	180	–
Sales (T)	–	–	6000
Maximum profit: \$ 59,950			

The mean profit is the average profit of the three scenarios which is \$ 115,406.

The problem for the farmer is that he has to decide on the land assignment, i.e., to determine x_1, x_2, x_3 without knowing which of the three scenarios is going to happen with regard to the purchases y_1, y_2 and sales w_1, w_2, w_3, w_4 which depend on the yield. The variables x_1, x_2, x_3 are called the first stage decision variables.

Hence, the decisions depend on the scenarios which are indexed by $j = 1, 2, 3$ with $j = 1$ referring to 'above average yields', $j = 2$ to 'average yields' and $j = 3$ to 'below average yields'. We introduce corresponding new variables $y_{ij}, 1 \le i \le 2, 1 \le j \le 3$, and $w_{ij}, 1 \le i \le 4, 1 \le j \le 3$. For instance, w_{31} represents the amount of sugar beets sold at the favorable price in case 'above average yields'. The decision variables y_{ij}, w_{ij} are referred to as the second stage decision variables.

We assume that the three scenarios occur at the same probability of 1/3. If the objective is to maximize long-run profit, we are led to the following problem:

$$
\begin{aligned}
(1.2) \quad & \text{minimize} \\
& \Big(150x_1 + 230x_2 + 260x_3 \\
& -\frac{1}{3}(170w_{11} - 238y_{11} + 150w_{21} - 210y_{21} + 36w_{31} + 10w_{41}) \\
& -\frac{1}{3}(170w_{12} - 238y_{12} + 150w_{22} - 210y_{22} + 36w_{32} + 10w_{42}) \\
& -\frac{1}{3}(170w_{13} - 238y_{13} + 150w_{23} - 210y_{23} + 36w_{33} + 10w_{43})\Big)
\end{aligned}
$$

$$
\begin{aligned}
\text{subject to} \quad & x_1 + x_2 + x_3 \le 500 , \\
& 3.0x_1 + y_{11} - w_{11} \ge 200 , \\
& 3.6x_2 + y_{21} - w_{21} \ge 240 ,
\end{aligned}
$$

$$
\begin{aligned}
w_{31} + w_{41} &\le 24x_3 ,\\
w_{31} &\le 6000 ,\\
2.5x_1 + y_{12} - w_{12} &\ge 200 ,\\
3.0x_2 + y_{22} - w_{22} &\ge 240 ,\\
w_{32} + w_{42} &\le 20x_3 ,\\
w_{32} &\le 6000 ,\\
2.0x_1 + y_{13} - w_{13} &\ge 200 ,\\
2.4x_2 + y_{23} - w_{23} &\ge 240 ,\\
w_{33} + w_{43} &\le 16x_3 ,\\
w_{33} &\le 6000 ,\\
x, y, w &\ge 0 .
\end{aligned}
$$

The optimization problem (1.2) is called a **stochastic decision problem**. In particular, (1.2) is said to be the **extensive form** of the stochastic program. The reason for this notation is that it explicitly describes the second stage variables for all possible scenarios. Its optimal solution is shown in Table 5.

TABLE 5. Solution of the stochastic decision problem (1.2)

		Wheat	Corn	Sugar Beets
First stage	Surface (acres)	170	80	250
$s = 1$	Yield (T)	510	288	6000
above	Purchases (T)	–	–	–
average	Sales (T)	310	48	6000
$s = 2$	Yield (T)	425	240	5000
average	Purchases (T)	–	–	–
	Sales (T)	225	–	5000
$s = 3$	Yield (T)	340	192	4000
below	Purchases (T)	–	48	–
average	Sales (T)	140	–	4000
Maximum profit: $ 108,390				

We see that the solution differs from those obtained in case of perfect a priori information. The distinctive feature is that in a stochastic setting the decisions have to be hedged against the various possible scenarios (cf. Tables 2,3 and 4). We also see that the **expected maximum profit** ($ 108,390) differs from the mean value ($ 115,406) of the maximum profits of the three scenarios in case of perfect a priori information. The difference \$ 7016 is called the **Expected Value of Perfect Information (EVPI)**.

A variant of the above stochastic decision problem is that the farmer makes the first stage decision (allocation of land) on the basis of 'average yields' according to Table 2. If the yields are again random with 20 % above resp. below average, it has to be observed that the planting costs are deterministic, but the purchases and sales depend on the yield. This leads to a reduced stochastic decision program where the maximum profit turns out to be \$ 107,240 which is in this case less than the maximum profit \$ 108,390 of the stochastic decision program (1.2). The difference \$ 1,150 is called the **Value of the Stochastic Solution (VSS)** reflecting the possible gain by solving the full stochastic model.

1.1.2 Two-stage stochastic program with recourse

For a stochastic decision program, we denote by $x \in \mathbb{R}^{n_1}, x \geq 0$, the vector of first stage decision variables. It is subject to constraints

$$Ax \ \leq \ b \ , \tag{1.3}$$

where $A \in \mathbb{R}^{m_1 \times n_1}, b \in \mathbb{R}^{m_1}$ are a fixed matrix and vector, respectively. In the optimal land usage problem, $x = (x_1, x_2, x_3)^T$ represents the amount of acres devoted to the three different crops. Here, $m_1 = 1$ and $A = (1\ 1\ 1), b = 500$.
We further denote by $\boldsymbol{\xi}$ a random vector whose realizations provide information on the second stage decisions $\mathbf{y}$ which is a random vector with realizations in $\mathbb{R}_+^{n_2}$. In the optimal land usage problem, $\boldsymbol{\xi} = (t_1, t_2, t_3)^T$ with $t_i = t_i(s), 1 \leq i \leq 3$, where $s \in \{1, 2, 3\}$ stands for the possible scenarios ('above average', 'average', and 'below average'). In other words, $t_i(s)$ represents the yield of crop i under scenario s.
$\mathbf{y} = (y_1, y_2, y_3 = w_1, y_4 = w_2, y_5 = w_3, y_6 = w_4)^T$ is the random vector whose realizations $\mathbf{y}(s), s \in \{1, 2, 3\}$, are the second stage decisions on the amount of crop to be purchased or sold at scenario s. The relationship between x and $\mathbf{y}$ can be expressed according to

$$W\mathbf{y} \ = \ \mathbf{h} \ - \ \mathbf{T}x \ , \tag{1.4}$$

where $\mathbf{h} \in \mathbb{R}^{m_2}$ is a fixed vector, $W \in \mathbb{R}^{m_2 \times n_2}$ is a fixed matrix, and $\mathbf{T}$ is a random matrix with realizations $\mathbf{T}(s) \in \mathbb{R}^{m_2 \times n_1}$.
W is called the **recourse matrix** and $\mathbf{T}$ is referred to as the **technology matrix**.

The **second stage decision problems** can be stated as

$$\text{(1.5)} \qquad \begin{aligned} &\text{minimize} \quad \mathbf{q}^T\mathbf{y} \\ &\text{subject to} \quad W\mathbf{y} \; + \; \mathbf{T}x \; \geq \; \mathbf{h}, \\ &\qquad\qquad\quad \mathbf{y} \; \geq \; 0 \end{aligned}$$

for given $\mathbf{q} \in \mathbb{R}^{n_2}$. We set

$$\text{(1.6)} \qquad Q(x,\boldsymbol{\xi}) \; := \; \min \; \{\mathbf{q}^T\mathbf{y} \mid W\mathbf{y} + \mathbf{T}x \; \geq \; \mathbf{h}\} \; .$$

In the optimal land usage problem, a second stage decision problem for scenario s can be written as

$$\text{(1.7)} \quad \begin{aligned} &\text{minimize} \; 238y_1 \; - \; 170w_1 \; + \; 210y_2 \; - \; 150w_2 \; - \; 36w_3 \; - \; 10w_4 \\ &\text{subject to} \quad t_1(s)x_1 \; + \; y_1 \; - \; w_1 \; \geq \; 200 \; , \\ &\qquad t_2(s)x_2 \; + \; y_2 \; - \; w_2 \; \geq \; 240 \; , \\ &\qquad t_3(s)x_3 \; - \; w_3 \; - \; w_4 \; \geq \; 0 \; , \\ &\qquad w_3 \; \leq \; 6000 \; , \\ &\qquad y, w \; \geq \; 0 \; . \end{aligned}$$

Altogether, the stochastic program can be formulated according to

$$\text{(1.8)} \qquad \begin{aligned} &\text{minimize} \; c^Tx \; + \; E_{\boldsymbol{\xi}}Q(x,\boldsymbol{\xi}) \\ &\text{subject to} \quad Ax \; \leq \; b \; , \\ &\qquad\qquad\quad x \geq \; 0 \; , \end{aligned}$$

where $c \in \mathbb{R}^{n_1}$ is given and $E_{\boldsymbol{\xi}}$ stands for the expectation.
The problem (1.8) is called a **two stage stochastic program with recourse**. It represents the implicit representation of the original stochastic decision problem (1.2).
Finally, we refer to the function

$$\text{(1.9)} \qquad \mathcal{Q}(x) \; := \; E_{\boldsymbol{\xi}}Q(x,\boldsymbol{\xi})$$

as the **value function** or **recourse function**. Then, in even more compact form (1.8) can be written as

$$\text{(1.10)} \qquad \begin{aligned} &\text{minimize} \; c^Tx \; + \; \mathcal{Q}(x) \\ &\text{subject to} \quad Ax \; \leq \; b \; , \\ &\qquad\qquad\quad x \geq \; 0 \; , \end{aligned}$$

1.1.3 Continuous random variables: The news vendor problem

As an example of a stochastic problem with continuous random variables we consider the so-called **news vendor problem**: The setting of the problem is as follows:
Every morning, a news vendor goes to the publisher and buys x newspapers at a price of c per paper. This number is bounded from above by x^{max}. The vendor tries to sell as many newspapers as possible at a selling price q. Any unsold newspapers can be returned to the publisher at a return price of $r < c$. The demand for newspapers varies of the days and is described by a continuous random variable $\boldsymbol{\xi}$ with probability distribution $F = F(\boldsymbol{\xi})$, i.e., $P(a \leq \boldsymbol{\xi} \leq b) = \int_a^b dF(\boldsymbol{\xi})$ and $\int_{-\infty}^{+\infty} dF(\boldsymbol{\xi}) = 1$.
The objective is to maximize the vendor's profit. To this end, we define y as the effective sales and w as the number of remittents. Then, the problem can be stated as

$$\begin{aligned} &\text{minimize } J(x) \ := \ cx \ + \ Q(x) \\ &\text{subject to } 0 \leq x \leq \ x^{max} \ , \end{aligned} \tag{1.11}$$

where

$$\begin{aligned} Q(x) \ &:= \ E_{\boldsymbol{\xi}} Q(x, \boldsymbol{\xi}) \ , \\ Q(x, \boldsymbol{\xi}) \ &:= \ \min \ - qy(\boldsymbol{\xi}) - rw(\boldsymbol{\xi}) \ , \\ &\quad \text{subject to } y(\boldsymbol{\xi}) \leq \boldsymbol{\xi} \ , \\ &\quad\quad y(\boldsymbol{\xi}) + w(\boldsymbol{\xi}) \leq x \ , \\ &\quad\quad y(\boldsymbol{\xi}), w(\boldsymbol{\xi}) \geq 0 \ . \end{aligned} \tag{1.12}$$

Note that $-Q(x)$ is the expected profit on sales and returns and $-Q(x, \boldsymbol{\xi})$ stands for the profit on sales and returns in case the demand is given by $\boldsymbol{\xi}$.
We see that like the optimal land usage problem, (1.11) represents another two-stage stochastic linear program with fixed recourse.
The optimal solution of (1.11) can be easily computed: When the demand $\boldsymbol{\xi}$ is known in the second stage, the optimal solution is given according to

$$y^*(\boldsymbol{\xi}) \ = \ \min(\boldsymbol{\xi}, x) \quad , \quad w^*(\boldsymbol{\xi}) \ = \ \max(x - \boldsymbol{\xi}, 0) \ ,$$

and hence, the second stage expected value function turns out to be

$$Q(x) \ = \ E_{\boldsymbol{\xi}} \left[-q \min(\boldsymbol{\xi}, x) \ - \ r \max(x - \boldsymbol{\xi}, 0) \right] \ .$$

The second stage expected value function can be computed by means of the probalibility distribution $F(\boldsymbol{\xi})$:

$$Q(x) = \int_{-\infty}^{x} \Big(- q\boldsymbol{\xi} - r(x - \boldsymbol{\xi})\Big) \, dF(\boldsymbol{\xi}) + \int_{x}^{+\infty} \Big(- qx\Big) \, dF(\boldsymbol{\xi}) = $$

$$= -(q-r) \int_{-\infty}^{x} \boldsymbol{\xi} \, dF(\boldsymbol{\xi}) - rxF(x) - qx(1 - F(x)) \ .$$

Integration by parts yields

$$\int_{-\infty}^{x} \boldsymbol{\xi} \, dF(\boldsymbol{\xi}) = xF(x) - \int_{-\infty}^{x} F(\boldsymbol{\xi}) \, d\boldsymbol{\xi} \ ,$$

whence

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

It follows that Q is differentiable in x with

$$Q'(x) = -q + (q-r)F(x) \ .$$

From Optimization I we know that the optimal solution x^* of (1.11) satisfies the variational equation

$$(v - x^*)J'(x^*) \geq 0 \ , \ v \in K := \{x \mid 0 \leq x \leq x^{max}\} \ ,$$

whose solution is given by

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

Since $J'(x) = c + Q'(x)$, we find

$$x^* = 0 \ , \ \text{if } \frac{q-c}{q-r} < F(0) \ ,$$

$$x^* = x^{max} \ , \ \text{if } \frac{q-c}{q-r} > F(x^{max}) \ ,$$

$$x^* = F^{-1}(\frac{q-c}{q-r}) \ , \ \text{otherwise} \ .$$

1.2 Two-stage stochastic linear programs with fixed recourse

1.2.1 Formulation of the problem and basic properties

Recalling the example concerning optimal land usage in case of stochastic uncertainties, we give the general formulation of a two-stage stochastic linear program with fixed recourse.

Definition 1.1 (Two-stage stochastic linear program with fixed recourse)
Let $I \subset \mathbb{N}$ be an index set, $A \subset \mathbb{R}^{m_1 \times n_1}$ and $W \in \mathbb{R}^{m_2 \times n_2}$ be fixed (deterministic) matrices, $T(\omega) \in \mathbb{R}^{m_2 \times n_1}, \omega \in I$, a random matrix, $c \in \mathbb{R}^{n_1}, b \in \mathbb{R}^{m_1}$ fixed (deterministic) vectors and $q(\omega) \in \mathbb{R}^{n_2}, h(\omega) \in \mathbb{R}^{m_2}, \omega \in I$, random vectors w.r.t. a probability space $(P, \Omega, \mathcal{A})$. Let further $\boldsymbol{\xi}^T(\omega) = (q(\omega)^T, h(\omega)^T, T_1(\omega), \cdots, T_{m_2}(\omega) \in \mathbb{R}^N, N := n_2 + m_2 + (m_2 \times n_1), \omega \in I$, where $T_i(\omega), 1 \leq i \leq m_2$, are the rows of $T(\omega)$. Then, the problem

$$\begin{aligned} \text{(1.13)} \qquad & \text{minimize } c^T x \; + \; E_{\boldsymbol{\xi}}(\min q(\omega)^T y(\omega)) \\ & Ax = b \ , \\ & T(\omega)x + Wy(\omega) = h(\omega) \text{ a.s. } , \\ & x \geq 0 \ , \ y(\omega) \geq 0 \text{ a.s. } , \end{aligned}$$

is called a **two-stage stochastic linear program with fixed recourse**. The matrix W is said to be the **recourse matrix** and the matrix $T(\omega)$ is referred to as the **technology matrix**.

Definition 1.2 (Deterministic Equivalent Program (DEP))
The linear program

$$\begin{aligned} \text{(1.14)} \qquad & \text{minimize } c^T x \; + \; Q(x) \\ & Ax = b \ , \\ & x \geq 0 \ , \end{aligned}$$

where

$$\begin{aligned} \text{(1.15)} \qquad Q(x) \; &:= \; E_{\boldsymbol{\xi}}(Q(x, \boldsymbol{\xi}(\omega))) \ , \\ Q(x, \boldsymbol{\xi}(\omega)) \; &:= \; \min_y \ \{q(\omega)^T y \mid Wy = h(\omega) - T(\omega)x \ , \ y \geq 0\} \ , \end{aligned}$$

is called the **Deterministic Equivalent Program (DEP)** associated with (1.13). The function Q is said to be the **recourse function** or **expected second-stage value function**.

Definition 1.3 (Feasible sets)
The sets

$$(1.16)\qquad \begin{aligned} K_1 &:= \{x \in \mathbb{R}^{n_1} \mid Ax = b\} , \\ K_2 &:= \{x \in \mathbb{R}^{n_1} \mid Q(x) < \infty\} \end{aligned}$$

are called the **first stage feasible set** and the **second stage feasible set**, respectively.
Let $\Sigma \subset \mathbb{R}^N$ be the support of $\boldsymbol{\xi}$ in the sense that $P(\{\boldsymbol{\xi} \in \Sigma\}) = 1$. If Σ is finite, $Q(x)$ is the weighted sum of finitely many $Q(x, \boldsymbol{\xi})$ values. We use the convention that if the values $\pm\infty$ occur, then $+\infty + (-\infty) = +\infty$. The sets

$$(1.17)\qquad \begin{aligned} K_2(\boldsymbol{\xi}) &:= \{x \in \mathbb{R}^{n_1} \mid Q(x, \boldsymbol{\xi}) < \infty\} , \\ K_2^P &:= \{x \in \mathbb{R}^{n_1} \mid \text{For all } \boldsymbol{\xi} \in \Sigma,\ y \geq 0 \text{ exists} \\ &\qquad \text{s.th. } Wy = h - Tx\} \ = \ \bigcap_{\boldsymbol{\xi} \in \Sigma} K_2(\boldsymbol{\xi}) \end{aligned}$$

are called the **elementary second stage feasible set** and the **possibility interpretation of the second stage feasible set**, respectively.

Definition 1.4 (Relatively complete, complete, and simple recourse)
The stochastic program (1.13) is said to have

- **relatively complete recourse**, if $K_1 \subset K_2$,
- **complete recourse**, if for all $z \in \mathbb{R}^{m_2}$ there exists $y \geq 0$ such that $Wy = z$,
- **simple recourse**, if the recourse matrix W has the structure $W = [I \ \ -I]$.

In case of simple recourse, we partition y and q according to $y = (y^+, y^-)$, and $q = (q^+, q^-)$. Then, the optimal values $(y_i^+(\omega), y_i^-(\omega)$ only depend on the sign of $h_i(\omega) - T_i(\omega)x$, provided $q_i = q_i^+ + q_i^- \geq 0$ with probability one. Moreover, if h_i has an associated distribution function F_i and mean value $\bar{h}_i$, there holds

$$(1.18)\quad Q_i(x) \ = \ q_i^+ \bar{h}_i \ - \ (q_i^+ - q_i F_i(T_i x)) T_i x \ - \ q_i \int\limits_{h_i \leq T_i x} h_i \, dF_i(h_i) \ .$$

Theorem 1.1 (Characterization of second stage feasible sets)
(i) For each $\boldsymbol{\xi}$, the elementary second stage feasible set $K_2(\boldsymbol{\xi})$ is a closed convex polyhedron which implies that K_2^P is a closed convex set.
(ii) Moreover, if Σ is finite, $K_2^P = K_2$.

Proof: The proof of (i) is obvious. For the proof of (ii) assume $x \in K_2$. Then, $Q(x)$ is bounded from above. Hence, $Q(x, \boldsymbol{\xi})$ is bounded from above for each $\boldsymbol{\xi}$ which shows $x \in K_2(\boldsymbol{\xi})$ for all $\boldsymbol{\xi}$, whence $x \in K_2^P$.
Conversely, assume $x \in K_2^P$. Then, $Q(x, \boldsymbol{\xi})$ is bounded from above for all $\boldsymbol{\xi}$. We deduce that $Q(x)$ is bounded from above and hence, $x \in K_2$.
□

We note that in case $\boldsymbol{\xi}$ is a continuous random variable similar results hold true, if $\boldsymbol{\xi}$ has finite second moments. For details we refer to [6] and [9].

Theorem 1.2 (Properties of the second stage value function)
Assume $Q(x, \boldsymbol{\xi}) > -\infty$. Then, there holds

(i) $Q(x, \boldsymbol{\xi})$ is piecewise linear and convex in (h, T).

(ii) $Q(x, \boldsymbol{\xi})$ is piecewise linear and concave in q.

(iii) $Q(x, \boldsymbol{\xi})$ is piecewise linear and convex in x for all $x \in K_1 \cap K_2$.

Proof: The piecewise linearity in (i)-(iii) follows from the existence of finitely many optimal bases for the second stage program. For details we refer to [7].
For the proof of the convexity in (h, t) resp. in x, it suffices to prove that the function

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

is convex in z. For $\lambda \in [0, 1]$ and $z_1, z_2, z_1 \neq z_2$, we consider $z(\lambda) := \lambda z_1 + (1 - \lambda) z_2$ and denote by $y_i^*, 1 \leq i \leq 2$, optimal solutions of the minimization problem for $z = z_1$ and $z = z_2$, respectively. Then, $y^*(\lambda) := \lambda y_1^* + [1 - \lambda] y_2^*$ is a feasible solution for $z = z(\lambda)$. If y_λ^* is the corresponding optimal solution, we obtain

$$\begin{aligned} g(z(\lambda)) \ &= \ q^T y_\lambda^* \ \leq \ q^T y^*(\lambda) \ = \\ &= \ \lambda q^T y_1^* \ + \ (1 - \lambda) q^T y_2^* \ = \ \lambda g(z_1) \ + \ (1 - \lambda) g(z_2) \ . \end{aligned}$$

The proof of the concavity in q is left as an exercise. □

For similar results in case $\boldsymbol{\xi}$ is a continuous random variable with finite second moments we again refer to [8].

1.2.2 Optimality conditions

For the derivation of the optimality conditions (KKT conditions), we assume that (1.14) has a finite optimal value. We refer to [8] for conditions that guarantee finiteness of the optimal value.

Theorem 1.3 (Optimality conditions)
Assume that (1.14) has a finite optimal value. A solution $x^* \in K_1$ of (1.14) is optimal if and only if there exist $\lambda^* \in \mathbb{R}^{m_1}, \mu^* \in \mathbb{R}^{n_1}_+, (\mu^*)^T x^* = 0$, such that

$$-c \ + \ A^T\lambda^* \ + \ \mu^* \in \partial Q(x^*) \ , \tag{1.19}$$

where $\partial Q(x^*)$ denotes the subdifferential of the recourse function Q.
Proof: As we know from Optimization I, the minimization problem

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

is a convex optimization problem with a closed convex constraint set which can be equivalently written as

$$\inf_{x \in \mathbb{R}^{n_1}} \ \sup_{\lambda \in \mathbb{R}^{m_1}, \mu \in \mathbb{R}^{n_1}_+} \ L(x, \lambda, \mu) \ , \tag{1.20}$$

where the Lagrangian is given by

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

The optimality condition for (1.20) reads

$$0 \in \partial L(x^*, \lambda^*, \mu^*) \ .$$

The subdifferential of the Lagrangian turns out to be

$$\partial L(x^*, \lambda^*, \mu^*) \ = \ c \ + \ \partial Q(x^*) \ - \ A^T\lambda^* \ - \ \mu^* \ ,$$

which results in (1.19). □

Obviously, the non-easy task will be to evaluate the subdifferential of the recourse function. The following result shows that it can be decomposed into subgradients of the recourse for each realization of $\boldsymbol{\xi}$.

Theorem 1.4 (Decomposition of the subgradient of the recourse function)
For $x \in K$ there holds

$$\begin{aligned} \partial Q(x) \ &= \ E_\omega \partial Q(x, \boldsymbol{\xi}(\omega)) \ + \ N(K_2, x) \ , \\ N(K_2, x) \ &= \ \{v \in \mathbb{R}^{n_1} \mid v^T y \leq 0 \text{ for all } y \text{ s.th. } x + y \in K_2\} \ , \end{aligned} \tag{1.21}$$

where $N(K_2, x)$ is the normal cone of the second stage feasible set K_2.
Proof: The subdifferential calculus of random convex functions with finite expectations [10] infers

$$\partial Q(x) \;=\; E_\omega \partial Q(x, \boldsymbol{\xi}(\omega)) \;+\; rec(\partial Q(x)) \;,$$

where $rec(\partial Q(x))$ is the recession cone of the subdifferential according to

$$rec(\partial Q(x)) \;=\; \{v \in \mathbb{R}^{n_1} \mid u + \lambda v \in \partial Q(x) \;,\; \lambda \geq 0 \;,\; u \in \partial Q(x)\} \;.$$

The recession cone can be equivalently written as

$$rec(\partial Q(x)) \;=\; \{v \in \mathbb{R}^{n_1} \mid y^T(u+\lambda v)+Q(x) \leq Q(x+y) \;,\; \lambda \geq 0 \;,\; y \in \mathbb{R}^{n_1}\} \;.$$

Consequently, we have

$$v \in rec(\partial Q(x)) \quad \Longleftrightarrow \quad y^T v \leq 0 \text{ for all } y \text{ s.th. } Q(x+y) < \infty \;.$$

Recalling the definition of K_2, we conclude. □

Corollary 1.5 (Optimality conditions in case of relatively complete recourse)
Assume that (1.14) has relatively complete recourse. Then, a solution $x^* \in K_1$ of (1.14) is optimal if and only if there exist $\lambda^* \in \mathbb{R}^{m_1}, \mu^* \in \mathbb{R}^{n_1}_+, (\mu^*)^T x^* = 0$, such that

$$-c \;+\; A^T\lambda^* \;+\; \mu^* \in E_\omega \partial Q(x^*, \boldsymbol{\xi}(\omega)) \;. \tag{1.22}$$

Proof: Taking into account that under the assumption of a relatively complete recourse there holds

$$N(K_2, x) \subset N(K_1, x) = \{v \in \mathbb{R}^{n_2} \mid v = A^T\lambda + \mu \;,\; \mu \geq 0 \;,\; \mu^T x = 0\} \;,$$

the result follows from Theorems 1.3 and 1.4. □

Corollary 1.6 (Optimality conditions in case of simple recourse)
Assume that (1.14) has relatively complete recourse. Then, a solution $x^* \in K_1$ of (1.14) is optimal if and only if there exist $\lambda^* \in \mathbb{R}^{m_1}, \mu^* \in \mathbb{R}^{n_1}_+, (\mu^*)^T x^* = 0$, and $\pi^* \in \mathbb{R}^{n_2}$ with

$$-(q_i^+ - q_i F_i(T_i x^*)) \;\leq\; \pi_i^* \;\leq\; -(q_i^+ - q_i F_i^+(T_i x^*))$$

where $F_i^+(h) := \lim_{t \to h_+} F_i(t)$, such that

$$-c \;+\; A^T\lambda^* \;+\; \mu^* \;-\; (\pi^*)^T T \;=\; 0 \;. \tag{1.23}$$

Proof: We deduce from (1.18) that

$$\partial Q_i(x) \;=\; \{\pi_i (T_i)^T \mid -(q_i^+ - q_i F_i(T_i x)) \;\leq\; \pi_i \;\leq\; -(q_i^+ - q_i F_i^+(T_i x))\} \;.$$

Then, (1.23) follows readily from Theorem 1.3. □

1.2.3 The value of information

Definition 1.5 (Expected value of perfect information)
For a particular realization $\boldsymbol{\xi} = \boldsymbol{\xi}(\omega), \omega \in I$, we consider the objective functional

$$J(x, \boldsymbol{\xi}) \ := \ c^T x \ + \ \min \ \{q^T y \mid Wy = h - Tx \ , \ y \geq 0\}$$

and the associated minimization problem

$$\min_{x \in K_1} E_{\boldsymbol{\xi}} J(x, \boldsymbol{\xi}) \quad , \quad K_1 \ := \ \{x \in \mathbb{R}_+^{n_1} \mid Ax = b\} \ .$$

The optimal solution is sometimes referred to as the **here-and-now solution** (cf. (1.13)). We denote the optimal value of this recourse problem by

$$RP \ := \ \min_{x \in K_1} J(x, \boldsymbol{\xi}) \ . \tag{1.24}$$

Another related minimization problem is to find the optimal solution for all possible scenarios and to consider the expected value of the associated optimal value

$$WS \ := \ E_{\boldsymbol{\xi}} \min_{x \in K_1} J(x, \boldsymbol{\xi}) \ . \tag{1.25}$$

The optimal solution of (1.25) is called the **wait-and-see solution**. The difference between the optimal values of the here-and-now solution and the wait-and-see solution

$$EVPI \ := \ RP \ - \ WS \tag{1.26}$$

is referred to as the **expected value of perfect information**.

Example: In the optimal land usage problem, we found

$$WS = -\$115,406 \quad , \quad RP = -\$108,390$$

so that $EVPI = \$7,016$. This amount represents the value of perfect information w.r.t. the wheather conditions for the next season.

The computation of the wait-and-see solution requires a considerable amount of computational work. Replacing all random variables by their expectations leads us to a quantity which is called the value of the stochastic solution.

Definition 1.6 (Value of the stochastic solution)
We denote by $\bar{\boldsymbol{\xi}} = E(\boldsymbol{\xi})$ the expectation of $\boldsymbol{\xi}$ and consider the minimization problem

$$EV \; := \; \min_{x \in K_1} J(x, \bar{\boldsymbol{\xi}}) \ . \tag{1.27}$$

which is dubbed the **expected value problem** or **mean value problem**. An optimal solution of (1.27) is called the **expected value solution**.
Denoting an optimal solution by $\bar{x}(\bar{\boldsymbol{\xi}})$, the expected value

$$EEV \; := \; E_{\boldsymbol{\xi}} J(\bar{x}(\bar{\boldsymbol{\xi}}), \boldsymbol{\xi}) \tag{1.28}$$

is referred to as the **expected result of using the expected value solution**. The difference

$$VSS \; := \; EEV \; - \; RP \tag{1.29}$$

is called the **value of the stochastic solution**. It measures the performance of $\bar{x}(\bar{\boldsymbol{\xi}})$ w.r.t. second stage decisions optimally chosen as functions of $\bar{x}(\bar{\boldsymbol{\xi}})$ and $\boldsymbol{\xi}$.

Example: In the optimal land usage problem, we have

$$EEV = -\$107,240 \quad , \quad RP = -\$108,390 \ ,$$

so that $VSS = \$1,150$. This amount represents the cost of ignoring uncertainty in the choice of a decision.

Theorem 1.7 (Fundamental inequalities, Part I)
Let RP , WS, and EEV be given by (1.24), (1.25) and (1.28), respectively. Then, there holds

$$WS \; \leq \; RP \; \leq \; EEV \ . \tag{1.30}$$

Proof: If x^* denotes the optimal solution of the recourse problem (1.24) and $\bar{x}(\boldsymbol{\xi})$ is the wait-and-see solution, we have

$$J(\bar{x}(\boldsymbol{\xi}), \boldsymbol{\xi}) \; \leq \; J(x^*, \boldsymbol{\xi}) \ .$$

Taking the expectation on both sides results in the left inequality in (1.30). Since x^* is the optimal solution of (1.24), whereas $\bar{x}(\bar{\boldsymbol{\xi}})$ is just one solution of the recourse problem, we arrive at the second inequality in (1.30). □

Theorem 1.8 (Fundamental inequalities, Part II)
Let WS and EV be given by (1.25) and (1.27). Then, in case of fixed (deterministic) objective coefficients and fixed (deterministic) technology matrix T, there holds

$$EV \ \leq \ WS \ . \tag{1.31}$$

Proof: We define

$$f(\boldsymbol{\xi}) \ = \ \min_{x \in K_1} J(x, \boldsymbol{\xi}) \ .$$

Recalling (1.25) and (1.27), we see that (1.31) is equivalent to

$$E(f(\boldsymbol{\xi})) \ \leq \ f(E(\boldsymbol{\xi})) \ . \tag{1.32}$$

Since (1.32) holds true for convex functions according to Jensen's inequality, the only thing we have to prove is the convexity of f. In order to do that, by duality we have

$$\min_{x \in K_1} J(x, \boldsymbol{\xi}) \ = \ \max_{\sigma, \pi} \ \{\sigma^T b + \pi^T h \mid \sigma^T A + \pi^T T \leq c^T \ , \ \pi^T W \leq q\} \ .$$

We observe that the constraints of the dual problem remain unchanged for all $\boldsymbol{\xi} = h$. Hence, epi f is the intersection of the epigraphs of the linear functions $\sigma^T b + \pi^T h$ for all feasible σ, π. The latter are obviously convex, and so is then epi f. We know from Optimization I that a function is convex if and only if its epigraph is convex. □

Theorem 1.9 (Fundamental inequalities, Part III)
Let RP and EEV be given by (1.24) and (1.28). Assume further that x^* is an optimal solution of (1.24) and that $\bar{x}(\bar{\boldsymbol{\xi}})$ is a solution of the expected value problem (1.28). Then, there holds

$$RP \ \geq \ EEV \ + \ (x^* - \bar{x}(\bar{\boldsymbol{\xi}}))^T \eta \ , \ \eta \in \partial E_{\boldsymbol{\xi}} J(\bar{x}(\bar{\boldsymbol{\xi}}), \boldsymbol{\xi}) \ . \tag{1.33}$$

Proof: The proof is left as an exercise. □

We finally derive an upper bound for the optimal value RP of the recourse problem (1.24) which is based on the observation

$$RP \ = \ \min_{x \in K_1} E_{\boldsymbol{\xi}} \bar{J}(x, \boldsymbol{\xi}) \ , \tag{1.34}$$

$$\bar{J}(x, \boldsymbol{\xi}) \ := \ c^T x \ + \ \min \ \{q^T y \mid Wy \geq h(\boldsymbol{\xi}) - Tx \ , \ y \geq 0\} \ . \tag{1.35}$$

Note that in (1.36) the second stage constraints are inequalities.

Theorem 1.10 (Fundamental inequalities, Part IV)
Assume that $h(\boldsymbol{\xi})$ is bounded from above, i.e., there exists h_{max} such that $h(\boldsymbol{\xi}) \leq h_{max}$ for all possible realizations of $\boldsymbol{\xi}$. Let x_{max} be an optimal solution of

$$\min_{x \in K_1} \bar{J}(x, h_{max}) \ .$$

Then, there holds

$$RP \ \leq \ \bar{J}(x_{max}, h_{max}) \ . \tag{1.36}$$

Proof: We see that for any $\boldsymbol{\xi} \in \Sigma$ and $x \in K_1$, a feasible solution of $Wy \geq h_{max} - Tx, y \geq 0$, is also a feasible solution of $Wy \geq h(\boldsymbol{\xi}) - Tx, y \geq 0$. Consequently, we have

$$\bar{J}(x, h_{max}) \ \geq \ \bar{J}(x, h(\boldsymbol{\xi})) \quad \Longrightarrow \quad \bar{J}(x, h_{max}) \ \geq \ E_{\boldsymbol{\xi}} \bar{J}(x, h(\boldsymbol{\xi})) \ ,$$

whence

$$\bar{J}(x, h_{max}) \ \geq \ \min_{x \in K_1} E_{\boldsymbol{\xi}} \bar{J}(x, h(\boldsymbol{\xi})) \ = \ RP \ . \quad \square$$

There is no universal relationship between $EVPI$ and VSS. For a discussion of this issue we refer to [1].

1.3 Numerical solution of two-stage stochastic linear programs with fixed recourse

1.3.1 The L-shaped method

We consider the deterministic equivalent program of a two-stage stochastic linear program with fixed recourse (cf. (1.13) and (1.14))

$$\begin{aligned} &\text{minimize } c^T x \;+\; \mathcal{Q}(x) \\ &\text{subject to } Ax = b \;,\; x \geq 0 \;, \end{aligned} \tag{1.37}$$

where

$$\begin{aligned} \mathcal{Q}(x) \;&:=\; E_{\boldsymbol{\xi}}(Q(x, \boldsymbol{\xi}(\omega))) \;, \\ Q(x, \boldsymbol{\xi}(\omega)) \;&:=\; \min_y \; \{q(\omega)^T y \mid Wy = h(\omega) - T(\omega)x \;,\; y \geq 0\} \;. \end{aligned} \tag{1.38}$$

The computational burden w.r.t. the DEP (1.37),(1.38) is the solution of all second stage recourse linear programs. If the random vector $\boldsymbol{\xi}$ only has a finite number, let's say, K realizations with probabilities $p_k, 1 \leq k \leq K$, the computational work can be significantly reduced by associating one set of second stage decisions y_k to each realization of $\boldsymbol{\xi}$, i.e., to each realization of q_k, h_k, and $T_k, 1 \leq k \leq K$. In other words, we consider the following **extensive form**

$$\begin{aligned} &\text{minimize } c^T x \;+\; \sum_{k=1}^{K} p_k q_k^T y_k \\ &\text{subject to } Ax = b \;, \\ &\qquad T_k x + W y_k \;=\; h_k \;,\; 1 \leq k \leq K \;, \\ &\qquad x \geq 0 \;,\; y_k \geq 0 \;,\; 1 \leq k \leq K \;. \end{aligned} \tag{1.39}$$

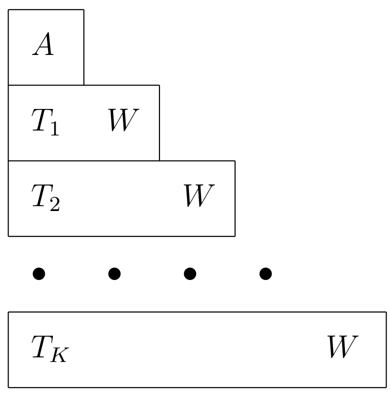

Fig.1.1. Block structure of the L-shaped method

The L-shaped method is an iterative process with **feasibility cuts** and **optimality cuts** according to the block structure of the extensive program illustrated in Fig. 1.1 which gives the method its name.

L-shaped algorithm

Step 0 (Initialization): Set $r = s = \nu = 0$ and $D_r = 0, d_r = 0, E_s = 0, e_s = 0$.

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

$$\text{(1.40a)} \qquad \text{minimize } J(x,\theta) \ := \ c^T x \ + \ \theta$$

$$\text{(1.40b)} \qquad \text{subject to } Ax = b \ ,$$

$$\text{(1.40c)} \qquad D_\ell \, x \ \geq \ d_\ell \ , \ \ell = 1, \cdots, r \ ,$$

$$\text{(1.40d)} \qquad E_\ell \, x + \theta \ \geq \ e_\ell \ , \ \ell = 1, \cdots, s \ ,$$

$$\text{(1.40e)} \qquad x \geq 0 \ , \ \theta \in \mathbb{R} \ .$$

If no constraints (1.40d) are present, set $\theta = 0$. Denote an optimal solution by (x^ν, θ^ν). Set $\theta^\nu = -\infty$, if there are no constraints (1.40d).

Step 2 (Feasibility cuts): For $k = 1, \cdots, K$ until $\tilde{J}(y, v^+, v^-) > 0$ solve the linear program

$$\text{(1.41a)} \qquad \text{minimize } \tilde{J}(y, v^+, v^-) \ := \ e^T v^+ \ + \ e^T v^-$$

$$\text{(1.41b)} \qquad \text{subject to } Wy \ + \ v^+ \ - \ v^- \ = \ h_k \ - \ T_k x^\nu \ ,$$

$$\text{(1.41c)} \qquad y \geq 0 \ , \ v^+ \geq 0 \ , \ v^- \geq 0 \ ,$$

where $e := (1, \cdots, 1)^T$. For the first $1 \leq k \leq K$ with $\tilde{J}(y, v^+, v^-) > 0$ let σ^ν be the associated Lagrange multiplier and define the **feasibility cut**

$$\text{(1.42a)} \qquad D_{r+1} \ := \ (\sigma^\nu)^T T_k \ ,$$

$$\text{(1.42b)} \qquad d_{r+1} \ := \ (\sigma^\nu)^T h_k \ . \ .$$

Set $r = r + 1$ and go back to Step 1.
If $\tilde{J}(y, v^+, v^-) = 0$ for all $1 \leq k \leq K$, go to Step 3.

Step 3 (Optimality cuts): For $k = 1, \cdots, K$ solve the linear program

$$\text{(1.43a)} \qquad \text{minimize } \hat{J}(y) \ := \ q^T y$$

$$\text{(1.43b)} \qquad \text{subject to } Wy \ = \ h_k \ - \ T_k x^\nu \ ,$$

$$\text{(1.43c)} \qquad y \ \geq \ 0 \ .$$

Let π_k^ν be the Lagrange multipliers associated with an optimal solution and define the **optimality cut**

$$E_{s+1} := \sum_{k=1}^K p_k(\pi_k^\nu)^T T_k \ , \tag{1.44a}$$

$$e_{s+1} := \sum_{k=1}^K p_k(\pi_k^\nu)^T h_k \ . \ . \tag{1.44b}$$

Set $J^\nu := e_{s+1} - E_{s+1}x^\nu$. If $\theta^\nu \geq J^\nu$, stop the iteration: x^ν is an optimal solution. Otherwise, set $s = s + 1$ and go back to Step 1.

1.3.2 Illustration of feasibility and optimality cuts

Example (Optimality cuts): We consider the minimization problem

$$\begin{aligned} &\text{minimize } Q(x) \\ &\text{subject to } 0 \leq x \leq 10 \ , \end{aligned} \tag{1.45}$$

where

$$Q(x,\boldsymbol{\xi}) := \begin{cases} \boldsymbol{\xi} - x \ , \ x \leq \boldsymbol{\xi} \\ x - \boldsymbol{\xi} \ , \ x \geq \boldsymbol{\xi} \end{cases} \ , \tag{1.46}$$

where $\boldsymbol{\xi}_1 = 1, \boldsymbol{\xi}_2 = 2, \boldsymbol{\xi}_3 = 4$ are the possible realizations of $\boldsymbol{\xi}$ with the probabilities $p_k = 1/3, 1 \leq k \leq 3$. Note that

$$W = 1 \ , \ T_k = \begin{cases} 1 \ , \ x \leq \boldsymbol{\xi}_k \\ -1 \ , \ x > \boldsymbol{\xi}_k \end{cases} \ , \ q_k = 1 \ , \ h_k = \begin{cases} \boldsymbol{\xi}_k \ , \ x \leq \boldsymbol{\xi}_k \\ -\boldsymbol{\xi}_k \ , \ x > \boldsymbol{\xi}_k \end{cases} \ .$$

Set $r = s = 0$, $D_r = 0, d_r = 0, E_s = 0, e_s = 0$, and $\nu = 1, x^1 = 0, \theta^1 = -\infty$, and begin the iteration with Step 2:
Iteration 1: In Step 2, we find $\tilde{J}(y, v^+, v^-) = 0, 1 \leq k \leq 3$, since $x^1 = 0$ is feasible. In Step 3, the solution of (1.43) yields $y = (1,2,4)^T$ with $\pi_k^1 = 1, 1 \leq k \leq 3$. Hence, (1.44a),(1.44b) give rise to

$$E_1 = 1 \quad , \quad e_1 = \frac{7}{3} \quad , \quad J^1 = \frac{7}{3} \ .$$

Set $s = 1$ and begin Iteration 2.
Iteration 2: In Step 1, the solution of the minimization problem

$$\begin{aligned} &\text{minimize } \theta \\ &\text{subject to } \theta \geq \frac{7}{3} - x \ , \\ &\qquad 0 \leq x \leq 10 \ , \ \theta \in \mathbb{R} \end{aligned}$$

is $x^2 = 10, \theta^2 = -\frac{23}{3}$.
Step 2 does not result in a feasibility cut, since x^2 is feasible. In Step

3, the solution of (1.43) yields $y = (9,8,6)^T$ with $\pi_k^2 = 1, 1 \leq k \leq 3$. Hence, from (1.44a),(1.44b) we obtain

$$E_2 = -1 \quad , \quad e_2 = -\frac{7}{3} \quad , \quad J^2 = \frac{23}{3} \ .$$

Set $s = 2$ and begin Iteration 3.

Iteration 3: In Step 1, the solution of the minimization problem

$$\begin{aligned} &\text{minimize } \theta \\ &\text{subject to } \theta \geq \frac{7}{3} - x \ , \\ &\qquad\qquad\quad \theta \geq x - \frac{7}{3} \ , \\ &\qquad\qquad\quad 0 \leq x \leq 10 \ , \ \theta \in \mathbb{R} \end{aligned}$$

is $x^3 = \frac{7}{3}, \theta^3 = 0$.
Step 2 does not result in a feasibility cut, since x^3 is feasible. In Step 3, the solution of (1.43) gives $y = (4/3, 1/3, 5/3)^T$ with $\pi_k^3 = 1, 1 \leq k \leq 3$. The equations (1.44a),(1.44b) imply

$$E_3 = -\frac{1}{3} \quad , \quad e_3 = \frac{1}{3} \quad , \quad J^3 = \frac{10}{9} \ .$$

Set $s = 3$ and begin Iteration 4.

Iteration 4: In Step 1, the solution of the minimization problem

$$\begin{aligned} &\text{minimize } \theta \\ &\text{subject to } \theta \geq \frac{7}{3} - x \ , \\ &\qquad\qquad\quad \theta \geq x - \frac{7}{3} \ , \\ &\qquad\qquad\quad \theta \geq \frac{x}{3} + \frac{1}{3} \ , \\ &\qquad\qquad\quad 0 \leq x \leq 10 \ , \ \theta \in \mathbb{R} \end{aligned}$$

is $x^4 = \frac{3}{2}, \theta^4 = \frac{5}{6}$.
Step 2 does not result in a feasibility cut, since x^4 is feasible. In Step 3, the solution of (1.43) results in $y = (1/2, 1/2, 5/2)^T$ with $\pi_k^4 = 1, 1 \leq k \leq 3$. The equations (1.44a),(1.44b) imply

$$E_4 = \frac{1}{3} \quad , \quad e_4 = \frac{5}{3} \quad , \quad J^4 = \frac{7}{6} \ .$$

Set $s = 4$ and begin Iteration 5.

Iteration 5: In Step 1, the solution of the minimization problem

$$\begin{aligned} &\text{minimize } \theta \\ &\text{subject to } \theta \ge \frac{7}{3} - x \ , \\ &\qquad \theta \ge x - \frac{7}{3} \ , \\ &\qquad \theta \ge \frac{x}{3} + \frac{1}{3} \ , \\ &\qquad \theta \ge \frac{5}{3} - \frac{x}{3} \ , \\ &\qquad 0 \le x \le 10 \ , \ \theta \in \mathbb{R} \end{aligned}$$

is $x^5 = 2, \theta^5 = 1$.
Step 2 does not result in a feasibility cut, since x^5 is feasible. In Step 3, the solution of (1.43) is $y = (1,0,2)^T$ with $\pi_k^\nu = 1, 1 \le k \le 3$. The equations (1.44a),(1.44b) yield

$$E_5 = \frac{1}{3} \quad , \quad e_5 = \frac{5}{3} \quad , \quad J^5 = 1 \ .$$

Since $J^5 = \theta^5$, we found the optimal solution.

Example (Feasibility cuts): We consider the minimization problem

$$\begin{aligned} &\text{minimize } 3x_1 + 2x_2 + E_{\boldsymbol{\xi}}(15y_1 + 12y_2) \\ &\text{subject to } 3y_1 + 2y_2 \le x_1 \ , \\ &\qquad 2y_1 + 5y_2 \le x_2 \ , \\ &\qquad 0.8\boldsymbol{\xi}_1 \le y_1 \le \boldsymbol{\xi}_1 \ , \\ &\qquad 0.8\boldsymbol{\xi}_2 \le y_2 \le \boldsymbol{\xi}_2 \ , \\ &\qquad x \ge 10 \ , \ y \ge 0 \ , \end{aligned} \tag{1.47}$$

where $\boldsymbol{\xi} = (\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)^T$ with $\boldsymbol{\xi}_1 \in \{4,6\}, \boldsymbol{\xi}_2 \in \{4,8\}$ independently with probability 1/2 each.
This example represents an investment decision in two resources x_1 and x_2 which are needed in the second stage decision to cover 80 % of the demand.
Note that

$$c = (3,2)^T \quad , \quad p_k = \frac{1}{2} \ , \ 1 \le k \le 2 \quad , \quad q = (15,12)^T \ ,$$
$$W = \begin{pmatrix} 3 & 2 \\ 2 & 5 \end{pmatrix} \quad , \quad T = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \ .$$

Consider the realization $\boldsymbol{\xi} = (6,8)^T$. Set $r = s = 0$, $D_r = 0, d_r = 0, E_s = 0, e_s = 0$, and $\nu = 1, x^1 = (0,0)^T, \theta^1 = -\infty$, and begin the iteration with Step 2 which results in a first feasibility cut

$$3x_1 \; + \; x_2 \; \geq \; 123.2 \; .$$

The associated first-stage solution provided by Step 1 is

$$x^1 \; = \; (41.067, 0)^T \; .$$

The following Step 2 gives rise to the feasibility cut

$$x_2 \; \geq \; 22.4 \; .$$

Going back to Step 1 and computing the associated first-stage solution gives

$$x^2 \; = \; (33.6, 22.4)^T \; .$$

Step 2 results in a third feasibility cut

$$x_2 \; \geq \; 41.6$$

with the associated first-stage solution

$$x^3 \; = \; (27.2, 41.6)^T$$

which guarantees feasible second-stage decisions.

Remark: This example illustrates that the formal application of the feasibility cuts can lead to an inefficient procedure. A closer look at the problem at hand shows that in case $\boldsymbol{\xi}_1 = 6$ and $\boldsymbol{\xi}_2 = 8$, feasibility requires

$$x_1 \; \geq \; 27.2 \quad , \quad x_2 \; \geq \; 41.6 \; .$$

In other words, a reasonable initial program is given by

$$\begin{aligned} &\text{minimize} \quad 3x_1 + 2x_2 + Q(x) \; , \\ &\text{subject to } x_1 \geq \; 27.2 \; , \\ &\qquad\qquad\quad x_2 \geq 41.6 \; , \end{aligned}$$

which guarantees second-stage feasibility.
Another particular case where second-stage feasibility is guaranteed (and thus Step 2 of the L-shaped method can be skipped) is a two-stage stochastic linear program with **complete recourse**, i.e., there exists $y \geq 0$ such that $Wy = t$ for all $t \in \mathbb{R}^{m_2}$.
For further specific cases where the structure of the program simplifies second-stage feasibility, we refer to [1].

Example: We illustrate the implementation of the L-shaped method for the following example:

$$n_1 = 1\ , n_2 = 6\ ,\ m_1 = 0\ ,\ m_2 = 3\ ,$$
$$c = 0\ ,\ W = \begin{pmatrix} 1 & -1 & -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}\ .$$

The random variable $\boldsymbol{\xi}$ has $K = 2$ independent realizations with probability 1/2 each which are given by

$$\boldsymbol{\xi}_1 = (q_1, h_1, T_1)^T \quad , \quad \boldsymbol{\xi}_2 = (q_2, h_2, T_2)^T\ ,$$

where

$$\begin{aligned} q_1 &= (1,0,0,0,0,0)^T \quad , \quad q_2 = (3/2, 0, 2/7, 1, 0, 0)^T\ , \\ h_1 &= (-1,2,7)^T \quad , \quad h_2 = (0,2,7)^T\ , \\ T_1 &= (1,0,0)^T \quad , \quad T_2 = T_1\ . \end{aligned}$$

We note that for $\boldsymbol{\xi}_1$, the recourse function is given by

$$Q_1(x) = \begin{cases} -x-1\ ,\ x \leq -1 \\ 0\ ,\ x \geq -1 \end{cases}\ ,$$

whereas for $\boldsymbol{\xi}_2$ we obtain

$$Q_2(x) = \begin{cases} -1.5x\ ,\ x \leq 0 \\ 0\ ,\ 0 \leq x \leq 2 \\ \frac{2}{7}(x-2)\ ,\ 2 \leq x \leq 9 \\ x-7\ ,\ x \geq 9 \end{cases}\ .$$

We further impose the constraints

$$-20 \ \leq\ x\ \leq\ +20\ .$$

A closer look at the problem reveals that $x = 0$ is an optimal solution.

Step 0 (Initialization): We choose $x^0 \leq -1$.

Iteration 1: We obtain

$$x^1 = -2\ ,\ \theta^1 \text{ is omitted}\ ,\ \text{New cut: } \theta \geq -0.5 - 1.25x\ .$$

Iteration 2: The second iteration yields

$$x^2 = 20\ ,\ \theta^2 = -25.5\ ,\ \text{New cut: } \theta \geq -3.5 + 0.5x\ .$$

Iteration 3: The computations result in

$$x^3 = \frac{12}{7}\ ,\ \theta^3 = -\frac{37}{14}\ ,\ \text{New cut: } \theta \geq 0\ .$$

Iteration 4: We get

$$x^4 \in [-\frac{2}{5}, 7] \ , \ \theta^4 = 0 \ .$$

If we choose $x^4 \in [0, 2]$, iteration 4 terminates. Otherwise, more iterations are needed.

1.3.3 Finite termination property

In this section, we prove that the L-shaped method terminates after a finite number of steps, provided $\boldsymbol{\xi}$ is a finite random variable. In particular, we show that

- a finite number of feasibility cuts (1.40b) is required either to provide a feasible vector within the second stage feasible set $K_2 = \{x \in \mathbb{R}^{n_1} \mid Q(x) < \infty\}$ (cf. (1.16)) or to detect infeasibility of the problem,
- a finite number of optimality cuts (1.40c) is needed to end up with an optimal solution of (1.37),(1.38), provided there exist feasible points $x \in K_2$.

We first recall the definition of $Q(x)$ in (1.37)

$$\begin{aligned} Q(x) &= E_\omega Q(x, \xi(\omega)) \ , \\ Q(x, \xi(\omega)) &= \min_y \ \{q(\omega)^T y \mid Wy = h(\omega) - T(\omega)x \ , \ y \geq 0\} \ . \end{aligned}$$

Lemma 1.11 (Representation of subgradients)
Let π_k^ν , $k \in \{1, \cdots, K\}$, be an optimal multiplier associated with an optimal solution x^ν of the minimization problem (1.43). Then, there holds

$$-(\pi_k^\nu)^T T_k \in \partial Q(x^\nu, \xi_k) \ . \tag{1.48}$$

Proof: The proof is left as an exercise. □

Proposition 1.12 (Finite termination of Step 2)
After a finite number of sub-steps, Step 2 (feasibility cuts) of the L-shaped method either terminates with feasible points $x \in K_2$ or detects infeasibility of (1.37),(1.38).

Proof: We introduce the set

$$\text{pos } W \ := \ \{t \in \mathbb{R}^{m_2} \mid t = Wy, y \geq 0\} \tag{1.49}$$

and note that

$$x \in K_2 \quad \Longleftrightarrow \quad h_k - T_k x \in \text{pos } W \ , \ 1 \leq k \leq K \ . \tag{1.50}$$

Given a first stage solution x^ν, in Step 2 the minimization problem (1.41) tests whether $h_k - T_k x^\nu \in \text{pos } W$ for all $1 \le k \le K$, or if $h_{\tilde{k}} - T_{\tilde{k}} x^\nu \notin \text{pos } W$ for some $\tilde{k} \in \{1, \cdots, K\}$. In the latter case, there exists a hyperplane separating $h_{\tilde{k}} - T_{\tilde{k}} x^\nu$ and pos W, i.e., there exists $\sigma \in \mathbb{R}^{m_2}$ such that

$$\sigma^T t \le 0 \ , \ t \in \text{pos } W \quad \text{and} \quad \sigma^T(h_{\tilde{k}} - T_{\tilde{k}} x^\nu) > 0$$

We remark that σ can be chosen as a multiplier $\sigma_{\tilde{k}}^\nu$ associated with (1.41), since

$$(\sigma_{\tilde{k}}^\nu)^T W \le 0 \quad \text{and} \quad (\sigma_{\tilde{k}}^\nu)^T (h_{\tilde{k}} - T_{\tilde{k}} x^\nu) > 0 \ .$$

On the other hand, we have

$$(1.51) \qquad x \in K_2 \quad \Longrightarrow \quad (\sigma_k^\nu)^T (h_k - T_k x) \le 0 \ , \ 1 \le k \le K \ .$$

Due to the finiteness of $\boldsymbol{\xi}$, there is only a finite number of optimal bases of problem (1.41) and hence, there are only finitely many constraints $(\sigma_k^\nu)^T(h_k - T_k x) \le 0$. Consequently, after a finite number of sub-steps we either find feasible points $x \in K_2$ or detect infeasibility. □

Proposition 1.13 (Finite termination of Step 3)
Assume feasibility of (1.37),(1.38). Then, after a finite number of sub-steps, Step 3 (optimality cuts) of the L-shaped method terminates with an optimal solution.

Proof: We first observe that (1.37),(1.38) can be equivalently stated as

$$(1.52) \qquad \begin{gathered} \text{minimize } c^T x \ + \ \Theta \\ Q(x) \le \Theta \ , \\ \text{subject to } x \in K_1 \cap K_2 \ , \end{gathered}$$

where K_1 is the first stage feasible set $K_1 := \{x \in \mathbb{R}^{n_1} \mid Ax = b\}$ (cf. (1.16)).
We further note that in Step 3 of the L-shaped method we compute the solution of (1.43) along with an associated multiplier π_k^ν. We know from the duality theory of linear programming (cf., e.g., Theorem 1.3(ii) in [4]) that

$$(1.53) \qquad Q(x^\nu, \boldsymbol{\xi}_k) \ = \ (\pi_k^\nu)^T (h_k - T_k x^\nu) \ , \ 1 \le k \le K \ .$$

Moreover, taking

$$v \in \partial Q(x^\nu, \boldsymbol{\xi}_k) \iff v^T(x - x^\nu) + Q(x^\nu, \boldsymbol{\xi}_k) \le Q(x, \boldsymbol{\xi}_k) \ , \ x \in \mathbb{R}^{n_1}$$

into account, it follows from (1.48) in Lemma 1.11 that

$$(1.54) \qquad (\pi_k^\nu)^T T_k (x^\nu - x) \ + \ Q(x^\nu, \boldsymbol{\xi}_k) \ \le \ Q(x, \boldsymbol{\xi}_k)$$

Using (1.53) in (1.54), we find

$$Q(x, \boldsymbol{\xi}_k) \;\geq\; (\pi_k^\nu)^T (h_k - T_k x) \; . \tag{1.55}$$

Denoting by $\mathbf{T}, \mathbf{h}$ and $\boldsymbol{\pi}^\nu$ the random variables with realizations T_k, h_k and $\pi_k^\nu, 1 \leq k \leq K$, respectively, and taking the expectations in (1.53) and (1.55), we get

$$Q(x^\nu) \;=\; E(\boldsymbol{\pi}^\nu)^T (\mathbf{h} - \mathbf{T} x^\nu) \;=\; \sum_{k=1}^K p_k (\pi_k^\nu)^T (h_k - T_k x^\nu)$$

and

$$Q(x) \;=\; E(\boldsymbol{\pi}^\nu)^T (\mathbf{h} - \mathbf{T} x) \;=\; \sum_{k=1}^K p_k (\pi_k^\nu)^T (h_k - T_k x) \; .$$

It follows that a pair (x, Θ) is feasible for (1.52) if and only if

$$\Theta \;\geq\; Q(x) \;\geq\; E(\boldsymbol{\pi}^\nu)^T (\mathbf{h} - \mathbf{T} x) \; ,$$

which corresponds to (1.40d).
On the other hand, if a pair (x^ν, Θ^ν) is optimal for (1.52), then

$$Q(x^\nu) \;=\; \Theta^\nu \;=\; E(\boldsymbol{\pi}^\nu)^T (\mathbf{h} - \mathbf{T} x^\nu) \; .$$

Consequently, at each sub-step of Step 3 we either find $\Theta^\nu \geq Q(x^\nu)$ which means that we found an optimal solution, or we have $\Theta^\nu < Q(x^\nu)$ which means that we have to continue with a new first stage solution $x^{\nu+1}$ and associated multipliers $\pi_k^{\nu+1}, 1 \leq k \leq K$ for (1.43). Since there is only a finite number of optimal bases associated with (1.43), there can be only a finite number of different combinations of the multipliers and hence, Step 3 must terminate after a finite number of sub-steps with an optimal solution for (1.37),(1.38). □

Unifying the results of Proposition 1.12 and Proposition 1.13, we arrive at the following finite convergence result:

Theorem 1.14 (Finite convergence of the L-shaped method) Assume that $\boldsymbol{\xi}$ is a finite random variable. Then, after a finite number of steps the L-shaped method either terminates with an optimal solution or proves infeasibility of (1.37),(1.38).

1.3.4 The multicut version of the L-shaped method

An alternative to Step 3 of the L-shaped method, where optimality cuts are computed with respect to the K realizations of the second-stage program and then aggregated to one cut (cf. (1.44a),(1.44b)), one can impose multiple cuts which leads to the following **multicut L-shaped algorithm:**

Multicut L-shaped algorithm

Step 0: Set $r = \nu = 0$ and $s_k = 0, 1 \le k \le K$.

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

$$\text{minimize } z(x) \ := \ c^T x \ + \ \sum_{k=1}^{K} \theta_k \ , \tag{1.56a}$$

$$\text{subject to } Ax = b \ , \tag{1.56b}$$

$$D_\ell \, x \ \ge \ d_\ell \ , \ \ell = 1, \cdots, r \ , \tag{1.56c}$$

$$E_{\ell(k)} \, x + \theta_k \ \ge \ e_{\ell(k)} \ , \ \ell(k) = 1, \cdots, s(k) \ , \quad 1 \le k \le K \ , \tag{1.56d}$$

$$x \ge 0 \ . \tag{1.56e}$$

Let $(x^\nu, \theta_1^\nu, \cdots, \theta_K^\nu)$ be an optimal solution of (1.56a)-(1.56e).
In case there are no constraints (1.56d) for some $k \in \{1, \cdots, K\}$, we set $\theta_k^\nu = -\infty$.

Step 2 (Feasibility cuts): Step 2 is performed as in Step of the L-shaped method.

Step 3 (Optimality cuts): For $1 \le kleK$ solve the linear programs (1.43) and denote by π_k^ν the optimal multiplier associated with the k-th problem. Check whether

$$\theta_k^\nu \ < \ p_k (\pi_k^\nu)^T (h_k - T_k x^\nu) \ . \tag{1.57}$$

If (1.57) is satisfied, define

$$E_{s(k)+1} \ = \ p_k (\pi_k^\nu)^T T_k \ , \tag{1.58a}$$

$$e_{s(k)+1} \ = \ p_k (\pi_k^\nu)^T h_k \ , \tag{1.58b}$$

set $s(k) = s(k) + 1$, and return to Step 1.
If (1.57) does not hold true for any $1 \le k \le K$, stop the algorithm: x^ν is an optimal solution.

Example: We consider the same example as in Chapter 1.3.2:

Step 0 (Initialization): We choose $x^0 \le -1$.

Iteration 1: We compute

$$x^1 = -2 \ , \ \theta_1^1 and \theta_2^1 \text{ are omitted} \ ,$$
$$\text{New cuts: } \theta_1 \geq -0.5 - 0.5x \ ,$$
$$\theta_2 \geq -\frac{3}{4}x \ .$$

Iteration 2: We obtain

$$x^2 = 20 \ , \ \theta_1^2 = -10.5 \ , \ \theta_2^2 = -15 \ ,$$
$$\text{New cuts: } \theta_1 \geq 0 \ ,$$
$$\theta_2 \geq -3.5 + 0.5x \ .$$

Iteration 3: The computations yield

$$x^3 = 2.8 \ , \ \theta_1^3 = 0 \ , \ \theta_2^3 = -2.1 \ ,$$
$$\text{New cut: } \theta_2 \geq \frac{1}{7}(x-2) \ .$$

Iteration 4: We get

$$x^4 = 0.32 \ , \ \theta_1^4 = 0 \ , \ \theta_2^4 = -0.24 \ ,$$
$$\text{New cut: } \theta_2 \geq 0 \ .$$

Iteration 5: This iteration reveals

$$x^5 = 0 \ , \ \theta_1^5 = 0 \ , \ \theta_2^5 = 0 \ .$$

The algorithm terminates with $x^5 = 0$ as an optimal solution.

1.3.5 Inner linearization methods

We consider the **dual linear program** with respect to (1.40a)-(1.40e) in Steps 1-3 of the L-shaped method:
Find (ρ, σ, π) such that

$$\text{(1.59a)} \quad \text{maximize } \zeta = \rho^T b + \sum_{\ell=1}^{r} \sigma_\ell d_\ell + \sum_{\ell=1}^{s} \pi_\ell e_\ell,$$

$$\text{(1.59b)} \quad \text{subj. to } \rho^T A + \sum_{\ell=1}^{r} \sigma_\ell D_\ell + \sum_{\ell=1}^{s} \pi_\ell E_\ell \ \leq \ c^T,$$

$$\text{(1.59c)} \quad \sum_{\ell=1}^{s} \pi_\ell = 1 \ , \ \sigma_\ell \geq 0, 0 \leq \ell \leq r \ , \ \pi_\ell \geq 0, 0 \leq \ell \leq s \ .$$

The dual program (1.59a)-(1.59c) involves

- **multipliers** $\sigma_\ell, 0 \leq \ell \leq r$, on **extreme rays** (directions of recession) of the duals of the subproblems,
- **multipliers** $\pi_\ell, 0 \leq \ell \leq s$, on the **expectations of extreme points** of the duals of the subproblems.

Indeed, let us consider the following dual linear program with respect to (1.43a)-(1.43c) in Step 3 (optimality cuts) of the L-shaped method:

$$\text{maximize } w \ = \ \pi^T(h_k - T_k x^\nu) \ , \tag{1.60a}$$

$$\text{subject to } \pi^T W \ \leq \ q^T \ . \tag{1.60b}$$

From the duality theory of linear programming we know (cf. Theorem 1.3 and Theorem 1.4 in [4]):

- If (1.60a)-(1.60b) is unbounded for all k, then there exists a multiplier σ^ν such that
  $$(\sigma^\nu)^T W \ \leq 0 \quad , \quad (\sigma^\nu)^T(h_k - T_k x^\nu) \ > \ 0 \ ,$$
  and the primal problem (1.41a)-(1.41c) does not have a feasible solution.
- If (1.60a)-(1.60b) is bounded for some k, then (1.60a)-(1.60b) is feasible and (1.41a)-(1.41c) has an optimal (primal) solution.

In other words, Step 2 of the L-shaped method is equivalent to checking whether (1.60a)-(1.60b) is unbounded for any k. If so, $D_{\ell+1}$ and d_ℓ are computed according to (1.42a) and (1.42b) of the L-shaped method and added to the constraints (feasibility cuts).

Next, consider the case when (1.60a)-(1.60b) has a finite optimal value σ_k^ν for all k, i.e., (1.41a)-(1.41c) is solvable for all k. In Step 3 of the L-shaped method, we then compute $E_{\ell+1}$ and $e_{\ell+1}$ according to (1.44a) and (1.44b) and add them to the constraints (optimality cuts). In the dual approach (1.59a)-(1.59c) we proceed in the same way.

Conclusion: Steps 1-3 of the L-shaped method are equivalent to solving (1.59a)-(1.59c) as a **master program** and the maximization problems (1.60a)-(1.60b) as **subproblems**.

This leads to the following so-called **inner linearization algorithm:**

Step 0: Set $r = s = \nu = 0$.

Step 1: Set $\nu = \nu + 1$. Compute $(\rho^\nu, \sigma^\nu, \pi^\nu)$ as the solution of (1.59a)-(1.59c) and (x^ν, θ^ν) as the associated dual solution.

Step 2: For $1 \leq k \leq K$ solve the subproblems (1.60a)-(1.60b).
If all subproblems are solvable, go to Step 3.
If an infeasible subproblem (1.60a)-(1.60b) is found, stop the algorithm

(the stochastic program is ill-posed).
If an unbounded solution with extreme ray σ^ν is found for some k, compute

$$D_{r+1} \ := \ (\sigma^\nu)^T T_k \ , \tag{1.61a}$$

$$d_{r+1} \ := \ (\sigma^\nu)^T h_k \ , \tag{1.61b}$$

set $r = r + 1$ and return to Step 1.

Step 3: Compute E_{s+1} and e_{s+1} according to

$$E_{s+1} \ := \ \sum_{k=1}^{K} p_k (\pi_k^\nu)^T T_k \ , \tag{1.62a}$$

$$e_{s+1} \ := \ \sum_{k=1}^{K} p_k (\pi_k^\nu)^T h_k \ , \tag{1.62b}$$

If

$$e_{s+1} \ - \ E_{s+1} x^\nu \ - \ \theta^\nu \ \leq \ 0 \ , \tag{1.63}$$

then stop: $(\rho^\nu, \sigma^\nu, \pi^\nu)$ and (x^ν, θ^ν) are optimal solutions.
On the other hand, if

$$e_{s+1} \ - \ E_{s+1} x^\nu \ - \ \theta^\nu \ > \ 0 \ , \tag{1.64}$$

set $s = s + 1$ and return to Step 1.

Remark: The name **inner linearization algorithm** stems from the fact that (1.59a)-(1.59c) can be interpreted as an inner linearization of the dual program of the original L-shaped method in the sense of the **Dantzig-Wolfe decomposition** of large-scale linear programs [3]. Since we solve dual problems instead of primal ones, finite convergence follows directly from the corresponding property of the original L-shaped method.

Remark: With regard to the **dimensionality** of the problems, in many applied cases we have $n_1 \gg m_1$. Then, the primal L-shaped method has basis matrices of order at most $m_1 + m_2$ compared to basis matrices of order $n_1 + n_1$ for the dual version. Therefore, the original (primal) L-shaped method is usually preferred.

The **inner linearization method** can be applied directly to the primal problem (1.14), if the technology matrix T is deterministic. In this case, (1.14) can be replaced by

$$\begin{aligned} \text{minimize } z &= c^T x + \mathbf{\Psi}(\chi) \\ & Ax = b \ , \\ & Tx - \chi = 0 \ , \\ & x \geq 0 \ , \end{aligned} \tag{1.65}$$

where

$$\begin{aligned} \mathbf{\Psi}(\chi) &:= E_{\boldsymbol{\xi}} \psi(\chi, \xi(\omega)) \ , \\ \psi(\chi, \xi(\omega)) &:= \min_{y \geq 0} \{ q(\omega)^T y \mid Wy = h(\omega) - \chi \} \ . \end{aligned} \tag{1.66}$$

The **idea** is to construct an inner linearization of the substitute $\mathbf{\Psi}(\chi)$ of the recourse function using the **generalized programming approach** from [2] by replacing $\mathbf{\Psi}(\chi)$ with the convex hull of points $\mathbf{\Psi}(\chi^\ell)$ computed within the iterations of the algorithm. In particular, each iteration generates an extreme point of a region of linearity for $\mathbf{\Psi}$.
We define $\mathbf{\Psi}_0^+(\zeta)$ as follows

$$\mathbf{\Psi}_0^+(\zeta) \ := \ \lim_{\alpha \to \infty} \frac{\mathbf{\Psi}(\chi + \alpha\zeta) - \mathbf{\Psi}(\chi)}{\alpha} \ . \tag{1.67}$$

Generalized programming algorithm for two-stage stochastic linear programs:

Step 0: Set $s = r = \nu = 0$.

Step 1: Set $\nu = \nu + 1$ and solve the **master linear program**

$$\text{minimize } z^\nu = c^T x + \sum_{i=1}^r \mu_i \mathbf{\Psi}_0^+(\zeta^i) + \sum_{i=1}^s \lambda_i \mathbf{\Psi}(\chi^i), \tag{1.68a}$$

$$\text{subj. to} \quad Ax = b \ , \tag{1.68b}$$

$$\text{subj. to} \quad Tx - \sum_{i=1}^r \mu_i \zeta^i - \sum_{i=1}^s \lambda_i \chi^i = 0, \tag{1.68c}$$

$$\sum_{\ell=1}^s \lambda_i = 1 \ , \ \lambda_i \geq 0, 1 \leq i \leq s \ , \tag{1.68d}$$

$$x \geq 0 \ , \ \mu_i \geq 0, 1 \leq i \leq r \ . \tag{1.68e}$$

If (1.68a)-(1.68e) is infeasible or unbounded, stop the algorithm.
Otherwise, compute the solution $(x^\nu, \mu^\nu, \lambda^\nu)$ and the dual solution $(\sigma^\nu, \pi^\nu, \rho^\nu)$.

Step 2: Solve the subproblem

$$\text{(1.69)} \qquad \text{minimize } \mathbf{\Psi}(\chi) \ + \ (\pi^\nu)^T\chi \ - \ \rho^\nu \quad \text{over } \chi \ .$$

If (1.69) has a solution χ^{s+1}, go to Step 3.
On the other hand, if (1.69) is unbounded, there exists a **recession direction** ζ^{r+1} such that for some χ

$$\mathbf{\Psi}(\chi + \alpha\zeta^{r+1}) \ + \ (\pi^\nu)^T(\chi + \alpha\zeta^{r+1}) \to -\infty \quad \text{as } \alpha \to +\infty \ .$$

In this case, we define

$$\text{(1.70)} \qquad \mathbf{\Psi}_0^+(\zeta^{r+1}) \ = \ \lim_{\alpha\to+\infty} \frac{\mathbf{\Psi}(\chi + \alpha\zeta) - \mathbf{\Psi}(\chi)}{\alpha} \ .$$

We set $r = r + 1$ and and return to Step 1.

Step 3: Check whether

$$\text{(1.71)} \qquad \mathbf{\Psi}(\chi^{s+1}) \ + \ (\pi^\nu)^T\chi^{s+1} \ - \ \rho^\nu \ \geq \ 0 \ .$$

If (1.71) holds true, stop the algorithm: $(x^\nu, \mu^\nu, \lambda^\nu)$ is an optimal solution of (1.65).
Otherwise, set $s = s + 1$ and return to Step 1.

Remark: In case of a **two-stage stochastic linear problem**, the subproblem (1.69) can be reformulated according to

$$\text{(1.72a)} \qquad \text{minimize } \sum_{k=1}^K p_k q_k^T y_k \ + \ (\pi^\nu)^T\chi \ - \ \rho^\nu \ ,$$

$$\text{(1.72b)} \qquad \text{subject to } \ Wy_k + \chi = h_k \quad , \quad 1 \leq k \leq K \ ,$$

$$\text{(1.72c)} \qquad \text{subject to } \ y_k \geq 0 \quad , \quad 1 \leq k \leq K \ .$$

In general, for $k \in \{1, \cdots, K\}$ the subproblem (1.72) can not be further separated into different subproblems so that the original L-shaped method should be preferred. However, for problems with a **simple recourse**, for each k the function $\mathbf{\Psi}(\chi)$ is separable into components, and (1.72) can be split into K independent subproblems.

Finite termination of the generalized programming algorithm will be established by means of the following result:

Proposition 1.15 (Characterization of extreme points)
Every optimal extreme point $(y_1^*, \cdots, y_K^*, \chi^*)$ of the feasible region of (1.72) corresponds to an extreme point χ^* of

$$\text{(1.73)} \qquad \{\chi \mid \mathbf{\Psi}(\chi) = (\pi^*)^T\chi + \theta\} \ ,$$

where $\pi^* = \sum_{k=1}^K \pi_k^*$ and each $\pi_k^*, 1 \le k \le K$, is an extreme point of

(1.74) $$\{\pi_k \mid \pi_k^T W \le q_k^T\} \ .$$

Proof. Let $(y_1^*, \cdots, y_K^*, \chi^*)$ be an optimal extreme point in (1.72). Then, we have

(1.75) $$q_k^T y_k^* \ \le \ q_k^T y_k \quad \text{for all } y_k \text{ with } W y_k = \xi_k - \chi^* \ .$$

We claim that

(1.76) $$y_k^* \quad \text{also is an extreme point of } \{y_k \mid W y_k = \xi_k - \chi^* \ , \ y_k \ge 0\} \ .$$

Indeed, if (1.76) is not true, we could choose y_k^* as the arithmetic mean of two distinct feasible y_i^1 and y_i^2.
It follows from (1.75) and (1.76) that y_k^* has a **complementary dual solution** π_k^*, i.e.,
(1.77)
$$\pi_k^* \quad \text{is an extreme point of } \{\pi_k \mid \pi_k^T W \le q_k^T\} \quad \text{and} \quad (q_k^T - (\pi_k^*)^T W) y_k^* = 0 \ .$$

The proof of the assertion will now be provided by a **contradiction argument**: We assume that $(y_1^*, \cdots, y_K^*, \chi^*)$ is not an extreme point of the linearity region

(1.78) $$\mathbf{\Psi}(\chi) = (\pi^*)^T \chi + \theta \quad , \quad \theta = \mathbf{\Psi}(\chi^*) - (\pi^*)^T \chi^* \quad , \quad \pi^* = \sum_{k=1}^K \pi_k^* \ .$$

Then, χ^* must be the convex combination of two $\chi^i, 1 \le i \le 2$,i.e., $\chi^* = \lambda \chi^1 + (1-\lambda)\chi^2, 0 < \lambda < 1$, where

$$\mathbf{\Psi}(\chi^i) = (\pi^*)^T \chi^i + \theta \quad , \quad 1 \le i \le 2 \ .$$

We also claim that
(1.79)
$$\mathbf{\Psi}(\chi^i) \ = \ \sum_{k=1}^K q_k^T y_k^i \ , \quad \text{where } q_k^T y_k^i \ = \ (\pi_k^*)^T (h_k - \chi^i) \ , \ 1 \le i \le 2 \ .$$

Indeed, if (1.79) does not hold true, due to the feasibility of π_k^* we would have

$$q_k^T y_k^i \ > \ (\pi_k^*)^T (h_k - \chi^i) \ , \ 1 \le i \le 2 \ ,$$

which would imply

$$\mathbf{\Psi}(\chi^i) \ > \ (\pi^*)^T \chi^i + \theta \ .$$

We remark that (1.79) also implies

(1.80) $$((\pi_k^*)^T W - q_k^T)(\lambda y_k^1 + (1-\lambda) y_k^2) \ = \ 0 \ ,$$

and hence, due to the fact that y_k^* is an extreme point of the feasible set of the k-th recourse problem,

$$y_k^* \;=\; \lambda y_k^1 + (1-\lambda) y_k^2 \; . \tag{1.81}$$

It follows from (1.81) that

$$(y_1^*, \cdots, y_k^*, \chi^*) \;=\; \lambda(y_1^1, \cdots, y_k^1, \chi^1) + (1-\lambda)(y_1^2, \cdots, y_k^2, \chi^2) \; ,$$

which contradicts that $(y_1^*, \cdots, y_k^*, \chi^*)$ is an extreme point. □

Proposition 1.16 (Characterization of extreme rays)
Any extreme ray associated with subproblem (1.72) is an extreme ray of a region of linearity of $\mathbf{\Psi}(\chi)$.

Proof. The proof is left as an exercise. □

Theorem 1.17 (Finite convergence of the generalized programming algorithm)
The application of the generalized programming algorithm to problem (1.65) with subproblem (1.72) converges after a finite number of steps.

Proof. Each solution of subproblem (1.72) generates a new linear region extreme value. For a new extreme ray ζ^{r+1} we have

$$\mathbf{\Psi}_0^+(\zeta^{r+1}) \;+\; (\pi^\nu)^T \zeta^{r+1} \;<\; 0 \; , \tag{1.82}$$

whereas

$$\mathbf{\Psi}_0^+(\zeta^i) \;+\; (\pi^\nu)^T \zeta^i \;<\; 0 \quad , \quad 1 \le i \le r \; . \tag{1.83}$$

As far as a new extreme point χ^{s+1} is concerned, such a point is added to the constraints only if

$$\mathbf{\Psi}(\chi^{s+1}) + (\pi^\nu)^T \chi^{s+1} - \rho^\nu \;<\; 0 \; , \tag{1.84}$$

whereas

$$\mathbf{\Psi}(\chi^i) + (\pi^\nu)^T \chi^i - \rho^\nu \;<\; 0 \quad , \quad 1 \le i \le s \; . \tag{1.85}$$

Since the number of regions satisfying (1.82)-(1.85) is finite and each region has a finite number of extreme rays and extreme points, the algorithm must terminate after a finite number of steps. □

1.4 Two-stage stochastic nonlinear programs with recourse

In this section, we consider a generalization of stochastic two-stage linear programs with recourse to problems involving nonlinear functions.

Definition 1.7 (Two-stage stochastic nonlinear program with recourse)
Let $I \subset \mathbb{N}$ be an index set, $f^1 : \mathbb{R}^{n_1} \to \mathbb{R}, g_i^1 : \mathbb{R}^{n_1} \to \mathbb{R}, 1 \leq i \leq m_1$ and $f^2(\cdot,\omega) : \mathbb{R}^{n_2} \to \mathbb{R}, g_i^2(\cdot,\omega) : \mathbb{R}^{n_2} \to \mathbb{R}, 1 \leq i \leq m_2, t_i^2(\cdot,\omega) : \mathbb{R}^{n_1} \to \mathbb{R}, 1 \leq i \leq m_2$, functions that are continuous for any fixed $\omega \in I$ and measurable in ω for any fixed first argument. Then, a minimization problem of the form

$$\begin{aligned} \text{minimize } & z = f^1(x) + \mathcal{Q}(x) \ , \\ \text{subject to } & g_i^1(x) \leq 0 \ , \ 1 \leq i \leq \bar{m}_1 \ , \\ & g_i^1(x) = 0 \ , \ \bar{m}_1 + 1 \leq i \leq m_1 \ , \end{aligned} \tag{1.86}$$

where $\mathcal{Q}(x) = E_\omega[Q(x,\omega)]$ and

$$\begin{aligned} & Q(x,\omega) = \inf f^2(y(\omega),\omega) \ , \\ \text{subject to } \quad & t_i^2(x,\omega) + g_i^2(y(\omega),\omega) \leq 0 \ , \ 1 \leq i \leq \bar{m}_2 \ , \\ & t_i^2(x,\omega) + g_i^2(y(\omega),\omega) = 0 \ , \ \bar{m}_2 + 1 \leq i \leq m_2 \ , \end{aligned} \tag{1.87}$$

is called a **two-stage stochastic nonlinear program with recourse function** $\mathcal{Q}(x)$.

Remark: We note that the assumptions in Definition 1.7 imply that $Q(x,\omega)$ is measurable in ω for all $x \in \mathbb{R}^{n_1}$ and hence, the recourse $\mathcal{Q}(x)$ is well defined.

Definition 1.8 (First and second stage feasible sets)
The set

$$K_1 := \{x \in \mathbb{R}^{n_1} \mid g_i^1(x) \leq 0 \ , \ 1 \leq i \leq \bar{m}_1 \ , \ g_i^1(x) = 0 \ , \ \bar{m}_1 + 1 \leq i \leq m_1\} \tag{1.88}$$

is called the **first-stage feasible set**, whereas the sets

$$\begin{aligned} K_2(\omega) \ := \ & \{x \in \mathbb{R}^{n_2} \mid \text{There exists } y(\omega) \text{ such that} \\ & t_i^2(x,\omega) + g_i^2(y(\omega),\omega) \leq 0 \ , \ 1 \leq i \leq \bar{m}_2 \ , \\ & t_i^2(x,\omega) + g_i^2(y(\omega),\omega) = 0 \ , \ \bar{m}_2 + 1 \leq i \leq m_2\} \ , \end{aligned} \tag{1.89}$$

$$K_2 \ := \ \{x \in \mathbb{R}^{n_2} \mid \mathcal{Q}(x) < \infty\}$$

are referred to as the **second-stage feasible sets**.

Remark: We note that unlike the situation in Chapter 1.2 we do not consider fixed recourse in order to keep an utmost amount of generality. We also remark that in case of fixed recourse the optimality conditions depend on the form of the objective and constraint functions anyway.

In order to ensure **necessary and sufficient optimality conditions** for (1.86),(1.87) we impose the following assumptions on the objective and constraint functions:

A1 (Convexity):

- The functions $f^1, g_i^1 : \mathbb{R}^{n_1} \to \mathbb{R}, 1 \le i \le \bar{m}_1$, are convex,
- The functions $g_i^1 : \mathbb{R}^{n_1} \to \mathbb{R}, \bar{m}_1 + 1 \le i \le m_1$, are affine,
- The functions $f^2(\cdot,\omega), g_i^2(\cdot,\omega) : \mathbb{R}^{n_2} \to \mathbb{R}, 1 \le i \le \bar{m}_2$, are convex for all $\omega \in I$,
- The functions $g_i^2(\cdot,\omega) : \mathbb{R}^{n_2} \to \mathbb{R}, \bar{m}_2 + 1 \le i \le m_2$, are affine for all $\omega \in I$,
- The functions $t_i^2(\cdot,\omega) : \mathbb{R}^{n_1} \to \mathbb{R}, 1 \le i \le \bar{m}_2$, are convex for all $\omega \in I$,
- The functions $t_i^2(\cdot,\omega) : \mathbb{R}^{n_1} \to \mathbb{R}, \bar{m}_2 + 1 \le i \le m_2$, are affine for all $\omega \in I$.

A2 (Slater Condition):

- If $Q(x) < \infty$, for almost all $\omega \in I$, there exists $y(\omega)$ such that

$$t_i^2(x,\omega) + \ g_i^2(y(\omega),\omega) \ < \ 0 \quad , \quad 1 \le i \le \bar{m}_2$$

and

$$t_i^2(x,\omega) + \ g_i^2(y(\omega),\omega) \ = \ 0 \quad , \quad \bar{m}_2 + 1 \le i \le m_2 \ .$$

Theorem 1.18 (Convexity of the recourse function)
Assume that assumptions **(A1)** and **(A2)** hold true. Then, the recourse function $Q(x,\omega)$ is a convex function of x for all $\omega \in I$.

Proof. Suppose that $y_i, 1 \le i \le 2$, are solutions of (1.87) with respect to $x_i, 1 \le i \le 2$, respectively. We have to show that for $\lambda \in [0,1]$

$$(1.90) \quad y = \lambda y_1 + (1-\lambda) y_2 \quad \text{solves (1.87) for } x := \lambda x_1 + (1-\lambda) x_2$$

as well. This is an easy consequence of the assumptions and left as an exercise. □

Theorem 1.19 (Lower semicontinuity of the recourse function)
If the second-stage feasible set $K_2(\omega)$ is bounded for all $\omega \in I$, then the recourse function $Q(\cdot,\omega)$ is lower semicontinuous for all $\omega \in I$.

Proof. We have to show that for any $\bar{x} \in \mathbb{R}^{n_1}$ and $\omega \in I$ we have

$$Q(\bar{x},\omega) \ \le \ \liminf_{x \to \bar{x}} Q(x,\omega) \ . \tag{1.91}$$

Suppose that $\{x^\nu\}_{\nu \in \mathbb{N}}$ is a sequence in $\mathbb{R}^{n_1}$ such that $x^\nu \to \bar{x}$ as $\nu \to \infty$. Without restriction of generality, we may assume that $Q(x^\nu,\omega) < \infty, \nu \in \mathbb{N}$, since otherwise we will find a subsequence $\mathbb{N}'$ with that property.
By our assumptions, we find $y^\nu(\omega), \nu \in \mathbb{N}$, such that

$$\begin{aligned} t_i^2(x^\nu,\omega) + g_i^2(y^\nu(\omega),\omega) \le 0 \quad &, \quad 1 \le i \le \bar{m}_2 \ , \\ t_i^2(x^\nu,\omega) + g_i^2(y^\nu(\omega),\omega) = 0 \quad &, \quad \bar{m}_2 + 1 \le i \le m_2 \ . \end{aligned}$$

The boundedness assumption and the continuity of the functions imply that the sequence $\{y^\nu(\omega)\}_{\nu \in \mathbb{N}}$ is bounded, and hence, there exist $\bar{y}(\Omega)$ and a subsequence $\mathbb{N}' \subset \mathbb{N}$, such that $y^\nu(\omega) \to \bar{y}(\omega)$ as $\nu \in \mathbb{N}' \to \infty$ and

$$\begin{aligned} t_i^2(\bar{x},\omega) + g_i^2(\bar{y}(\omega),\omega) \le 0 \quad &, \quad 1 \le i \le \bar{m}_2 \ , \\ t_i^2(\bar{x},\omega) + g_i^2(\bar{y}(\omega),\omega) = 0 \quad &, \quad \bar{m}_2 + 1 \le i \le m_2 \ . \end{aligned}$$

Consequently, $\bar{x}$ is feasible and

$$Q(\bar{x},\omega) \ \le \ f^2(\bar{x},\omega) \ = \ \lim_{\nu \to \infty} f^2(x^\nu,\omega) \ = \ \lim_{\nu \to \infty} Q(x^\nu,\omega) \ ,$$

which gives the assertion. □

Corollary 1.20 (Further properties of the feasible set and the recourse function)
The feasible set K_2 is a closed, convex set, and the expected recourse function Q is a lower semicontinuous convex function in x.

Proof. The proof is an immediate consequence of the assumptions and the previous results. □

Remark: In general, it is difficult to decompose the feasible set K_2 according to

$$K_2 \ = \ \bigcap_{\omega \in I} K_2(\omega) \ . \tag{1.92}$$

A particular example, where such a decomposition can be realized is for quadratic objective functionals f^2.

Theorem 1.21 (Optimality conditions)
Suppose that there exists

$$x \in \mathrm{ri}(\mathrm{dom}(f^1(x))) \cap \mathrm{ri}(\mathrm{dom}(Q(x)) \ , \tag{1.93}$$

where ri stands for the relative interior, and further assume that

(1.94) $$g_i^1(x) < 0 \quad , \quad 1 \le i \le \bar{m}_1 \ ,$$

(1.95) $$g_i^1(x) = 0 \quad , \quad \bar{m}_1 + 1 \le i \le m_1 \ .$$

Then, $x^* \in \mathbb{R}^{n_1}$ is optimal in (1.86) if and only if $x^* \in K_1$ and there exist multipliers $\mu_i^* \ge 0, 1 \le i \le \bar{m}_1$, and $\lambda_i^*, \bar{m}_1 + 1 \le i \le m_1$, such that

(1.96) $$0 \in \partial f^1(x^*) + \partial Q(x^*) + \sum_{i=1}^{\bar{m}_1} \mu_i^* \partial g_i^1(x^*) + \sum_{i=\bar{m}_1+1}^{m_1} \lambda_i^* \partial g_i^1(x^*) \ ,$$

(1.97) $$\mu_i^* g_i^1(x^*) = 0 \quad , \quad 1 \le i \le \bar{m}_1 \ .$$

Proof. The assertions can be deduced readily by applying the general theory of nonlinear programming (cf., e.g., Chapter 2 in [4]). □

Remark: As far as decompositions of the subgradient $\partial Q(x)$ into subgradients of $Q(x, \omega)$ are concerned, in much the same way as in Theorem 1.4 of Chapter 1.2 one can show

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

where $N(K_2, x)$ stands for the normal cone (cf. Chapter 2 in [4]).
Note that (1.98) reduces to

$$\partial Q(x) \ = \ E_\omega[\partial Q(x, \omega)]$$

in case of relatively complete recourse, i.e., if $K_1 \subset K_2$.
For the derivation of optimality conditions for problems with explicit constraints on non-anticipativity we refer to Theorem 39 in [1].

1.5 Piecewise quadratic form of the L-shaped method

Although such a systematic approach like SQP (Sequential Quadratic Programming) for deterministic nonlinear problems does not exist in a stochastic environment, one might be tempted to reduce a general two-stage stochastic nonlinear program to the successive solution of two-stage stochastic quadratic programs. In this section, we consider a piecewise quadratic form of the L-shaped method for such two-stage stochastic quadratic programs which are of the form:

$$(1.99)\quad \text{minimize } z(x) = c^T x + \frac{1}{2} x^T C x \ + \ + \ E_{\boldsymbol{\xi}}[\min(q^T(\omega) y(\omega) + \frac{1}{2} y^T(\omega) D(\omega) y(\omega))] \ ,$$

$$\text{subject to} \quad Ax = b \ , \quad T(\omega)x + Wy(\omega) = h(\omega) \ , \quad x \geq 0 \ , \ y(\omega) \geq 0 \ .$$

Here, $A \in \mathbb{R}^{m_1 \times n_1}, C \in \mathbb{R}^{n_1 \times n_1}, W \in \mathbb{R}^{m_2 \times n_2}$ are fixed matrices, and $c \in \mathbb{R}^{n_1}$ is a fixed vector. Moreover, $D \in \mathbb{R}^{n_2 \times n_2}, T \in \mathbb{R}^{m_2 \times n_1}$ are random matrices and $q \in \mathbb{R}^{n_2}, h \in \mathbb{R}^{m_2}$ are random vectors.

For a given realization $\boldsymbol{\xi}(\omega), \omega \in I$, the associated recourse function can be defined according to

$$(1.100)\quad Q(x, \boldsymbol{\xi}(\omega)) \ := \ \min\{q^T(\omega) y(\omega) + \frac{1}{2} y^T(\omega) D(\omega) y(\omega) \ | \ T(\omega)x + Wy(\omega) = h(\omega) \ , \ y(\omega) \geq 0\} \ ,$$

which may attain the values $\pm\infty$, if the problem is unbounded or infeasible, respectively.
The expected recourse function is given by

$$(1.101)\qquad \mathcal{Q}(x) \ := \ E_{\boldsymbol{\xi}} Q(x, \boldsymbol{\xi}) \ .$$

We use the convention $+\infty + (-\infty) = +\infty$.
The first-stage and second-stage feasible sets K_1 and K_2 are defined as in the previous section.

We impose the following assumptions on the data of the two-stage stochastic quadratic program:

A3:

- The random vector $\boldsymbol{\xi}$ has a discrete distribution.

A4:

- The matrix C is positive semi-definite and the matrices $D(\omega)$ are positive semi-definite for all $\omega \in I$,
- The matrix W has full row rank.

Remark: We note that **(A13)** implies the decomposability of the second-stage feasible set K_2, whereas **(A4)** ensures convexity of the recourse functions.

An important feature of the problem is that the recourse function $Q(x)$ is **piecewise quadratic**, i.e., the second-stage feasible set K_2 can be decomposed into polyhedral sets, called **cells**, such that $Q(x)$ is quadratic on each cell.

Example: We consider the following two-stage stochastic quadratic program:

$$\begin{aligned} \text{(1.102)} \qquad & \text{minimize } z(x) = 2x_1 + 3x_2 + E_{\boldsymbol{\xi}} \min\{-6.5y_1 - \\ & \qquad - 7y_2 + \frac{1}{2}y_1^2 + y_1y_2 + \frac{1}{2}y_2^2\} \ , \\ & \text{subject to} \quad 3x_1 + 2x_2 \le 15 \ , \\ & \qquad x_1 + 2x_2 \le 8 \ , \\ & \qquad y_1 \le x_1 \ , \ y_2 \le x_2 \ , \\ & \qquad y_1 \le \boldsymbol{\xi}_1 \ , \ y_2 \le \boldsymbol{\xi}_2 \ , \\ & \qquad x_1 + x_2 \ge 0 \ , \ x \ge 0 \ , \ y \ge 0 \ . \end{aligned}$$

We assume that $\boldsymbol{\xi}_1 \in \{2,4,6\}$ and $\boldsymbol{\xi}_2 \in \{1,3,5\}$ are independent random variables with probability 1/3 for each realization.

The problem can be interpreted as a **portfolio problem** where the issue is to minimize quadratic penalties on deviations from a mean value.

In the second stage of the problem, for small values of the assets $x_i, 1 \le i \le 2$, it is optimal to sell, i.e., $y_i = x_i, 1 \le i \le 2$. Indeed, for

$$(x_1,x_2) \in C_1 := \{(x_1,x_2) \mid 0 \le x_1 \le 2 \ , \ 0 \le x_2 \le 1\}$$

the optimal solution of the second stage is $y_i = x_i, 1 \le i \le 2$ for all possible values of $\boldsymbol{\xi}$, whence

$$Q(x) = Q(x,\boldsymbol{\xi}) = -6.5x_1 - 7x_2 + \frac{1}{2}x_1^2 + x_1x_2 + \frac{1}{2}x_2^2 \ , \ (x_1,x_2) \in C_1 \ .$$

Definition 1.9 (Finite closed convex complex)
A **finite closed convex complex** $\mathcal{K}$ is a finite collection of closed convex sets $C_\nu, 1 \leq \nu \leq M$, called the **cells** of $\mathcal{K}$, such that

$$\mathrm{int}(C_{\nu_1} \cap C_{\nu_2}) = \emptyset, \nu_1 \neq \nu_2 \ .$$

Definition 1.10 (Piecewise convex program)
A **piecewise convex program** is a convex program of the form

$$\inf \{z(x) \mid x \in S\} \ , \tag{1.103}$$

where $z : \mathbb{R}^n \to \mathbb{R}$ is convex and S is a closed convex subset of $\mathrm{dom}(z)$ with $\mathrm{int}(S) \neq \emptyset$.

Definition 1.11 (Piecewise quadratic function)
Consider a piecewise convex program and assume that $\mathcal{K}$ is a finite closed convex complex with cells $C_\nu, 1 \leq \nu \leq M$ such that

(1.104a)

$$S \subseteq \bigcup_{\nu=1}^{M} C_\nu \ ,$$

(1.104b)

either $z \equiv -\infty$, or for each cell $C_\nu \ , \ 1 \leq \nu \leq M,$ there exists a convex function $z_\nu : S \to \mathbb{R}$ which is continuously differentiable on an open set containing C_ν such that

$$z(x) \ = \ z_\nu(x) \ , \ x \in C_\nu \ , \ 1 \leq \nu \leq M \ ,$$

$$\nabla z_\nu(x) \in \partial z(x) \ , \ x \in C_\nu \ , \ 1 \leq \nu \leq M \ .$$

A **piecewise quadratic function** $z : S \to \mathbb{R}$ is a piecewise convex function where on each cell $C_\nu, 1 \leq \nu \leq M$, the function z_ν is a quadratic form.

Example: In the above example, both $Q(x)$ and $z(x)$ are piecewise quadratic. In particular, we have

$$\begin{aligned} Q(x) &= -6.5x_1 - 7x_2 + \frac{1}{2}x_1^2 + x_1x_2 + \frac{1}{2}x_2^2 \ , \ (x_1, x_2) \in C_1 \ , \\ z(x) &= -4.5x_1 - 4x_2 + \frac{1}{2}x_1^2 + x_1x_2 + \frac{1}{2}x_2^2 \ , \ (x_1, x_2) \in C_1 \ . \end{aligned}$$

The **numerical solution** of two-stage stochastic piecewise quadratic programs is taken care of by the following **PQP algorithm:**

Step 0 (Initialization): Compute a decomposition of the state space S according to (1.104a) into cells $C_\nu, 1 \leq \nu \leq M$, set $S_1 = S$ and choose $x^0 \in S_1$.

Step 1 (Iteration loop): For $\mu \geq 1$:

Step 1.1 (Determination of current cell): Determine C_μ such that $x^{\mu-1} \in C_\mu$ and specify the quadratic form $z_\mu(\cdot)$ on C_μ according to (1.104b).

Step 1.2 (Solution of minimization subproblems): Compute

$$x^\mu = \arg\min_{x \in S_\mu} z_\mu(x) \ , \tag{1.105a}$$

$$w^\mu = \arg\min_{x \in C_\mu} z_\mu(x) \ . \tag{1.105b}$$

If w^μ is the limiting point of a ray on which $z_\mu(\cdot)$ is decreasing to $-\infty$, stop the algorithm: The original PQP is unbounded. Otherwise, continue with Step 1.3.

Step 1.3 (Optimality check): Check the optimality condition

$$(\nabla z_\mu(w^\mu))^T (x^\mu - w^\mu) = 0 \ . \tag{1.106}$$

If (1.106) is satisfied, stop the algorithm: w^μ is the optimal solution of the PQP. Otherwise, continue with Step 1.4.

Step 1.4 (Update of state space): Compute

$$S_{\mu+1} := S_\mu \cap \{x \mid (\nabla z_\mu(w^\mu))^T x \leq (\nabla z_\mu(w^\mu))^T w^\mu\} \ , \tag{1.107}$$

set $\mu := \mu + 1$, and go to Step 1.1.

Theorem 1.22 (Finite termination of the PQP algorithm)
Under assumptions **(A3)** and **(A4)**, the PQP algorithm terminates after a finite number of steps with the solution of the two-stage stochastic piecewise quadratic program.

Proof. We refer to [5]. □

Remark: Details concerning the appropriate construction of finite closed convex complexes $\mathcal{K}$ satisfying (1.104a),(1.104b) can be found in [5].

Example: We illustrate the **implementation of the PQP algorithm** for the piecewise quadratic program (1.102).

Step 0 (Initialization): We choose the cells $C_\nu, 1 \leq \nu \leq 8$, as shown in Fig. 1.2.
We further define

$$S_1 \;=\; S \;=\; \{x \in \mathbb{R}^2 \mid 3x_1 + 2x_2 \leq 15 \;,\; x_1 + 2x_2 \leq 8 \;,\; x_1, x_2 \geq 0\}$$

choose

$$x^0 \;=\; (0,0)^T$$

and set $\mu = 1$.

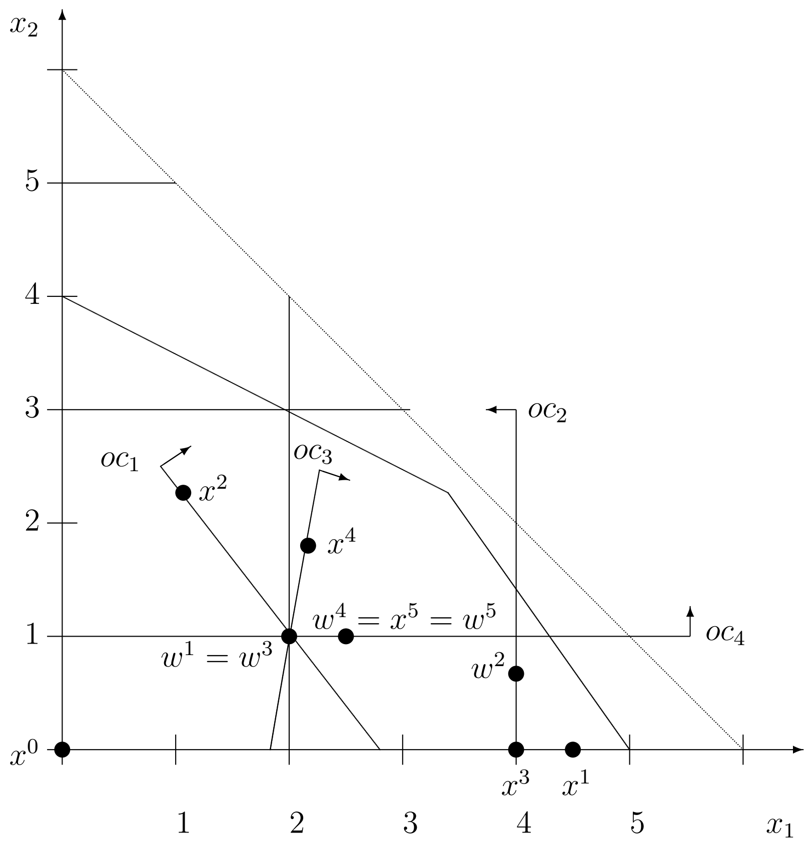

Fig. 1.2. Finite closed convex complex and PQP cuts

Iteration 1: The cell containing x^0 is

$$C_1 \ = \ \{x \in \mathbb{R}^2 \mid 0 \leq x_1 \leq 2 \ , \ 0 \leq x_2 \leq 1\} \ ,$$

and the quadratic function z_1 on C_1 is

$$z_1(x) \ = \ -4.5x_1 - 4x_2 + \frac{1}{2}x_1^2 + x_1x_2 + \frac{1}{2}x_2^2 \ .$$

Solving (1.105a),(1.105b) by means of the KKT-conditions results in

$$x^1 \ = \ (4.5, 0)^T \quad , \quad w^1 \ = \ (2, 1)^T \in C_1 \ ,$$

whence

$$\nabla z_1(w^1) \ = \ (-1.5, -1)^T \quad , \quad (\nabla z_1(w^1))^T(x^1 - w^1) \ = \ -2.75 \ \neq \ 0 \ ,$$

and

$$S_2 \ = \ S_1 \cap \{x \in \mathbb{R}^2 \mid -1.5x_1 - x_2 \leq -4\} \ .$$

Iteration 2: The cell containing x^1 is

$$C_2 \ = \ \{x \in \mathbb{R}^2 \mid 4 \leq x_1 \leq 6 \ , \ 0 \leq x_2 \leq 1 \ , \ x_1 + x_2 \leq 6.5\} \ .$$

The quadratic function z_2 on C_2 is

$$z_2(x) \ = \ -\frac{29}{3} - \frac{1}{6}x_1 - 2x_2 + \frac{1}{6}x_1^2 + \frac{1}{3}x_1x_2 + \frac{1}{2}x_2^2 \ .$$

The solution of (1.105a),(1.105b) by means of the KKT-conditions gives

$$x^2 \ = \ (\frac{22}{19}, \frac{43}{19})^T \quad , \quad w^2 \ = \ (4, \frac{2}{3})^T \in C_2 \ ,$$

whence

$$\nabla z_2(w^2) \ = \ (\frac{25}{18}, 0)^T \quad , \quad (\nabla z_2(w^2))^T(x^2 - w^2) \ \neq \ 0 \ ,$$

and

$$S_3 \ = \ S_2 \cap \{x \in \mathbb{R}^2 \mid \frac{25}{18}x_1 \leq \frac{100}{18}\} \ .$$

Iteration 3: The cell containing x^2 is

$$C_3 \ = \ \{x \in \mathbb{R}^2 \mid 0 \leq x_1 \leq 2 \ , \ 1 \leq x_2 \leq 3\} \ .$$

The quadratic function z_3 on C_3 is

$$z_3(x) \ = \ -\frac{13}{6} - \frac{25}{6}x_1 - \frac{5}{3}x_2 + 2x_1x_2 + \frac{1}{3}x_2^2 \ .$$

Via the solution of (1.105a),(1.105b) we obtain

$$x^3 \ = \ (4, 0)^T \quad , \quad w^3 = w^1 \ = \ (2, 1)^T \ ,$$

whence

$$S_4 \;=\; S_3 \cap \{x \in \mathbb{R}^2 \mid -\frac{3}{2}x_1 + \frac{1}{3}x_2 \le -\frac{8}{3}\} \; .$$

Iteration 4: The cell containing x^3 is

$$C_4 \;=\; \{x \in \mathbb{R}^2 \mid 2 \le x_1 \le 4 \; , \; 0 \le x_2 \le 1\} \; .$$

The quadratic function z_4 on C_4 is

$$z_4(x) \;=\; -\frac{11}{3} - \frac{7}{3}x_1 - \frac{10}{3}x_2 + \frac{1}{3}x_1^2 + \frac{2}{3}x_1x_2 + \frac{1}{3}x_2^2 \; .$$

The solution of (1.105a),(1.105b) yields

$$x^4 \;\approx\; (2.18, 1.81)^T \quad , \quad w^4 \;=\; (2.5, 1)^T \; ,$$

whence

$$S_5 \;=\; S_4 \cap \{x \in \mathbb{R}^2 \mid -\frac{2}{3}x_1 \le -\frac{2}{3}\} \; .$$

Iteration 5: The cell containing x^4 is

$$C_5 \;=\; \{x \in \mathbb{R}^2 \mid 2 \le x_1 \le 4 \; , \; 1 \le x_2 \le 3\} \cap S \; .$$

The quadratic function z_5 on C_5 is

$$z_4(x) \;=\; -\frac{101}{18} - \frac{19}{9}x_1 - \frac{11}{9}x_2 + \frac{1}{3}x_1^2 + \frac{4}{9}x_1x_2 + \frac{1}{3}x_2^2 \; .$$

The solution of (1.105a),(1.105b) yields

$$x^5 \;=\; w^5 \;=\; (2.5, 1)^T \; ,$$

which is an optimal solution of the problem.

References

[1] J.R. Birge and F. Louveaux; Introduction to Stochastic Programming. Springer, Berlin-Heidelberg-New York, 1997

[2] G.B. Dantzig; Linear Programming and Extensions. Princeton University Press, Princeton, NJ, 1963

[3] G.B. Dantzig and P. Wolfe; The decomposition principle for linear programs. Operations Research, **8**, 101–111, 1960

[4] R.H.W. Hoppe; Optimization I. Handout of the course held in Fall 2006. See http://www.math.uh.edu

[5] F.V. Louveaux; Piecewise convex programs. Math. Prgramming, **15**, 53–62, 1978

[6] D. Walkup and R.J.-B. Wets; Stochastic programs with recourse. SIAM J. Appl. Math., **15**, 1299–1314, 1967

[7] D. Walkup and R.J.-B. Wets; Stochastic programs with recourse II: on the continuity of the objective. SIAM J. Appl. Math., **17**, 98–103, 1969

[8] R.J.-B. Wets; Characterization theorems for stochastic programs. Math. Programming, **2**, 166–175, 1972

[9] R.J.-B. Wets; Stochastic programs with fixed recourse: the equivalent deterministic problem. SIAM Rev. **16**, 309–339, 1974

[10] R.J.-B. Wets; Stochastic programming. In: Optimization (G.L. Nemhauser et al.; eds.), Handbboks in Operations Research and Management Science, Vol. I, North-Holland, Amsterdam, 1990