AIMMS Modeling Guide - Linear Programming Tricks

AIMMS 4

This documentation was typeset by AIMMS B.V. using LATEX and the LUCIDA font family.

Part II

General Optimization Modeling Tricks

Chapter 6

Linear Programming Tricks

This chapter

This chapter explains several *tricks* that help to transform some models with special, for instance nonlinear, features into conventional linear programming models. Since the fastest and most powerful solution methods are those for linear programming models, it is often advisable to use this format instead of solving a nonlinear or integer programming model where possible.

References

The linear programming tricks in this chapter are not discussed in any particular reference, but are scattered throughout the literature. Several tricks can be found in [Wi90]. Other tricks are referenced directly.

Statement of a linear program

Throughout this chapter the following general statement of a linear programming model is used:

$$\begin{array}{lll} \textbf{Minimize:} & \displaystyle\sum_{j \in J} c_j x_j & \\ \textbf{Subject to:} & & \\ & \displaystyle\sum_{j \in J} a_{ij} x_j \gtrless b_i & \forall i \in I \\ & x_j \geq 0 & \forall j \in J \end{array}$$

In this statement, the c_j's are referred to as *cost coefficients*, the a_{ij}'s are referred to as *constraint coefficients*, and the b_i's are referred to as *requirements*. The symbol "$\gtrless$" denotes any of "$\leq$" , "$=$", or "$\geq$" constraints. A maximization model can always be written as a minimization model by multiplying the objective by (-1) and minimizing it.

6.1 Absolute values

The model

Consider the following model statement:

$$\begin{array}{lll} \textbf{Minimize:} & \displaystyle\sum_{j \in J} c_j |x_j| & c_j > 0 \\ \textbf{Subject to:} & & \\ & \displaystyle\sum_{j \in J} a_{ij} x_j \gtrless b_i & \forall i \in I \\ & x_j \quad \textit{free} & \end{array}$$

Instead of the standard cost function, a weighted sum of the absolute values of the variables is to be minimized. To begin with, a method to remove these absolute values is explained, and then an application of such a model is given.

Handling absolute values . . .

The presence of absolute values in the objective function means it is not possible to directly apply linear programming. The absolute values can be avoided by replacing each x_j and $|x_j|$ as follows.

$$\begin{aligned} x_j &= x_j^+ - x_j^- \\ |x_j| &= x_j^+ + x_j^- \\ x_j^+, x_j^- &\geq 0 \end{aligned}$$

The linear program of the previous paragraph can then be rewritten as follows.

Minimize: $$\sum_{j \in J} c_j(x_j^+ + x_j^-) \qquad c_j > 0$$

Subject to:

$$\sum_{j \in J} a_{ij}(x_j^+ - x_j^-) \gtreqless b_i \qquad \forall i \in I$$

$$x_j^+, x_j^- \geq 0 \qquad \forall j \in J$$

. . . correctly

The optimal solutions of both linear programs are the same if, for each j, at least one of the values x_j^+ and x_j^- is zero. In that case, $x_j = x_j^+$ when $x_j \geq 0$, and $x_j = -x_j^-$ when $x_j \leq 0$. Assume for a moment that the optimal values of x_j^+ and x_j^- are both positive for a particular j, and let $\delta = \min\{x_j^+, x_j^-\}$. Subtracting $\delta > 0$ from both x_j^+ and x_j^- leaves the value of $x_j = x_j^+ - x_j^-$ unchanged, but reduces the value of $|x_j| = x_j^+ + x_j^-$ by 2δ. This contradicts the optimality assumption, because the objective function value can be reduced by $2\delta c_j$.

Application: curve fitting

Sometimes x_j represents a *deviation* between the left- and the right-hand side of a constraint, such as in *regression*. Regression is a well-known statistical method of fitting a curve through observed data. One speaks of *linear regression* when a straight line is fitted.

Example

Consider fitting a straight line through the points (v_j, w_j) in Figure 6.1. The coefficients a and b of the straight line $w = av + b$ are to be determined.
The coefficient a is the *slope* of the line, and b is the *intercept* with the w-axis. In general, these coefficients can be determined using a model of the following form:

Minimize: $$f(z)$$

Subject to:

$$w_j = av_j + b - z_j \qquad \forall j \in J$$

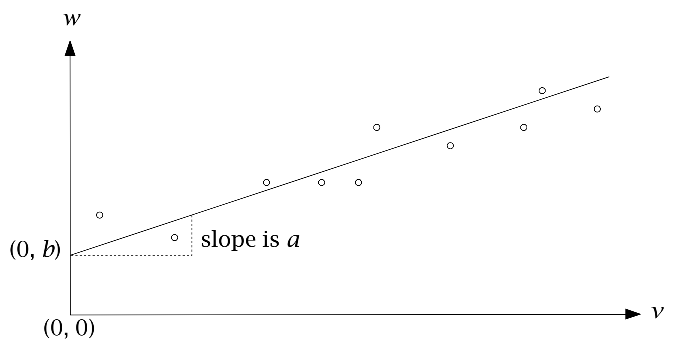


Figure 6.1: Linear regression

In this model z_j denotes the difference between the value of $av_j + b$ proposed by the linear expression and the observed value, w_j. In other words, z_j is the *error* or deviation in the w direction. Note that in this case a, b, and z_j are the decision variables, whereas v_j and w_j are data. A function $f(z)$ of the error variables must be minimized. There are different options for the objective function $f(z)$.

Different objectives in curve fitting

Least-squares estimation is an often used technique that fits a line such that the sum of the squared errors is minimized. The formula for the objective function is:

$$f(z) = \sum_{j \in J} z_j^2$$

It is apparent that quadratic programming must be used for least squares estimation since the objective is quadratic.

Least absolute deviations estimation is an alternative technique that minimizes the sum of the absolute errors. The objective function takes the form:

$$f(z) = \sum_{j \in J} |z_j|$$

When the data contains a few extreme observations, w_j, this objective is appropriate, because it is less influenced by extreme outliers than is least-squares estimation.

Least maximum deviation estimation is a third technique that minimizes the maximum error. This has an objective of the form:

$$f(z) = \max_{j \in J} |z_j|$$

This form can also be translated into a linear programming model, as explained in the next section.

6.2 A minimax objective

The model

Consider the model

Minimize: $\max_{k\in K} \sum_{j\in J} c_{kj}x_j$

Subject to:

$$\sum_{j\in J} a_{ij}x_j \gtrless b_i \qquad \forall i \in I$$
$$x_j \geq 0 \qquad \forall j \in J$$

Such an objective, which requires a maximum to be minimized, is known as a *minimax objective*. For example, when $K = \{1,2,3\}$ and $J = \{1,2\}$, then the objective is:

Minimize: $\max\{c_{11}x_1 + c_{12}x_2\ c_{21}x_1 + c_{22}x_2\ c_{31}x_1 + c_{32}x_2\}$

An example of such a problem is in least maximum deviation regression, explained in the previous section.

Transforming a minimax objective

The minimax objective can be transformed by including an additional decision variable z, which represents the maximum costs:

$$z = \max_{k\in K} \sum_{j\in J} c_{kj}x_j$$

In order to establish this relationship, the following extra constraints must be imposed:

$$\sum_{j\in J} c_{kj}x_j \leq z \qquad \forall k \in K$$

Now when z is minimized, these constraints ensure that z will be greater than, or equal to, $\sum_{j\in J} c_{kj}x_j$ for all k. At the same time, the optimal value of z will be *no greater than* the maximum of all $\sum_{j\in J} c_{kj}x_j$ because z has been minimized. Therefore the optimal value of z will be both as small as possible and exactly equal to the maximum cost over K.

The equivalent linear program

Minimize: z

Subject to:

$$\sum_{j\in J} a_{ij}x_j \gtrless b_i \qquad \forall i \in I$$
$$\sum_{j\in J} c_{kj}x_j \leq z \qquad \forall k \in K$$
$$x_j \geq 0 \qquad \forall j \in J$$

The problem of maximizing a minimum (a *maximin* objective) can be transformed in a similar fashion.

6.3 A fractional objective

The model

Consider the following model:

$$\begin{aligned}
\textbf{Minimize:} \quad & \Big(\sum_{j\in J} c_j x_j + \alpha\Big) \Big/ \Big(\sum_{j\in J} d_j x_j + \beta\Big) \\
\textbf{Subject to:} \quad & \\
& \sum_{j\in J} a_{ij} x_j \gtrless b_i \qquad \forall i \in I \\
& x_j \geq 0 \qquad \forall j \in J
\end{aligned}$$

In this problem the objective is the ratio of two linear terms. It is assumed that the denominator (the expression $\sum_{j\in J} d_j x_j + \beta$) is either positive or negative over the entire feasible set of x_j. The constraints are linear, so that a linear program will be obtained if the objective can be transformed to a linear function. Such problems typically arise in financial planning models. Possible objectives include the rate of return, turnover ratios, accounting ratios and productivity ratios.

Transforming a fractional objective

The following method for transforming the above model into a regular linear programming model is from Charnes and Cooper ([Ch62]). The main trick is to introduce variables y_j and t which satisfy: $y_j = t x_j$. In the explanation below, it is assumed that the value of the denominator is positive. If it is negative, the directions in the inequalities must be reversed.

1. Rewrite the objective function in terms of t, where

$$t = 1/(\sum_{j\in J} d_j x_j + \beta)$$

   and add this equality and the constraint $t > 0$ to the model. This gives:

$$\begin{aligned}
\textbf{Minimize:} \quad & \sum_{j\in J} c_j x_j t + \alpha t \\
\textbf{Subject to:} \quad & \\
& \sum_{j\in J} a_{ij} x_j \gtrless b_i \qquad \forall i \in I \\
& \sum_{j\in J} d_j x_j t + \beta t = 1 \\
& t > 0 \\
& x_j \geq 0 \qquad \forall j \in J
\end{aligned}$$

2. Multiply both sides of the original constraints by t, ($t > 0$), and rewrite the model in terms of y_j and t, where $y_j = x_j t$. This yields the model:

$$\begin{aligned}
&\textbf{Minimize:} && \sum_{j\in J} c_j y_j + \alpha t \\
&\textbf{Subject to:} && \\
&&& \sum_{j\in J} a_{ij} y_j \gtrless b_i t \qquad \forall i \in I \\
&&& \sum_{j\in J} d_j y_j + \beta t = 1 \\
&&& t > 0 \\
&&& y_j \geq 0 \qquad \forall j \in J
\end{aligned}$$

3. Finally, temporarily allow t to be ≥ 0 instead of $t > 0$ in order to get a linear programming model. This linear programming model is equivalent to the fractional objective model stated above, provided $t > 0$ at the optimal solution. The values of the variables x_j in the optimal solution of the fractional objective model are obtained by dividing the optimal y_j by the optimal t.

6.4 A range constraint

The model

Consider the following model:

$$\begin{aligned}
&\textbf{Minimize:} && \sum_{j\in J} c_j x_j \\
&\textbf{Subject to:} && \\
&&& d_i \leq \sum_{j\in J} a_{ij} x_j \leq e_i \qquad \forall i \in I \\
&&& x_j \geq 0 \qquad \forall j \in J
\end{aligned}$$

When one of the constraints has both an upper and lower bound, it is called a *range constraint*. Such a constraint occurs, for instance, when a minimum amount of a nutrient is required in a blend and, at the same time, there is a limited amount available.

Handling a range constraint

The most obvious way to model such a range constraint is to replace it by two constraints:

$$\begin{aligned}
&\sum_{j\in J} a_{ij} x_j \geq d_i \qquad \text{and} \\
&\sum_{j\in J} a_{ij} x_j \leq e_i \qquad \forall i \in I
\end{aligned}$$

However, as each constraint is now stated twice, both must be modified when changes occur. A more elegant way is to introduce extra variables. By introducing new variables u_i one can rewrite the constraints as follows:

$$u_i + \sum_{j\in J} a_{ij} x_j = e_i \quad \forall i \in I$$

The following bound is then imposed on u_i:

$$0 \leq u_i \leq e_i - d_i \quad \forall i \in I$$

It is clear that $u_i = 0$ results in

$$\sum_{j \in J} a_{ij} x_j = e_i$$

while $u_i = e_i - d_i$ results in

$$\sum_{j \in J} a_{ij} x_j = d_i$$

The equivalent linear program

A summary of the formulation is:

Minimize:

$$\sum_{j \in J} c_j x_j$$

Subject to:

$$\begin{aligned} u_i + \sum_{j \in J} a_{ij} x_j &= e_i & \forall i \in I \\ x_j &\geq 0 & \forall j \in J \\ 0 \leq u_i &\leq e_i - d_i & \forall i \in I \end{aligned}$$

6.5 A constraint with unknown-but-bounded coefficients

This section

This section considers the situation in which the coefficients of a linear inequality constraint are unknown-but-bounded. Such an inequality in terms of uncertainty intervals is not a deterministic linear programming constraint. Any particular selection of values for these uncertain coefficients results in an unreliable formulation. In this section it will be shown how to transform the original nondeterministic inequality into a set of deterministic linear programming constraints.

Unknown-but-bounded coefficients

Consider the constraint with unknown-but-bounded coefficients $\tilde{a}_j$

$$\sum_{j \in J} \tilde{a}_j x_j \leq b$$

where $\tilde{a}_j$ assumes an unknown value in the interval $[L_j, U_j]$, b is the fixed right-hand side, and x_j refers to the solution variables to be determined. Without loss of generality, the corresponding bounded uncertainty intervals can be written as $[a_j - \Delta_j, a_j + \Delta_j]$, where a_j is the midpoint of $[L_j, U_j]$.

Midpoints can be unreliable

Replacing the unknown coefficients by their midpoint results in a deterministic linear programming constraint that is not necessarily a reliable representation of the original nondeterministic inequality. Consider the simple linear program

$$\begin{array}{ll}
\textbf{Maximize:} & x \\
\textbf{Subject to:} & \\
 & \tilde{a}x \leq 8
\end{array}$$

with the uncertainty interval $\tilde{a} \in [1, 3]$. Using the midpoint $a = 2$ gives the optimal solution $x = 4$. However, if the true value of $\tilde{a}$ had been 3 instead of the midpoint value 2, then for $x = 4$ the constraint would have been violated by 50%.

Worst-case analysis

Consider a set of arbitrary but fixed x_j values. The requirement that the constraint with unknown-but-bounded coefficients must hold for the unknown values of $\tilde{a}_j$ is certainly satisfied when the constraint holds for all possible values of $\tilde{a}_j$ in the interval $[a_j - \Delta_j, a_j + \Delta_j]$. In that case it suffices to consider only those values of $\tilde{a}_j$ for which the term $\tilde{a}_j x_j$ attains its maximum value. Note that this situation occurs when $\tilde{a}_j$ is at one of its bounds. The sign of x_j determines which bound needs to be selected.

$$\begin{aligned}
\tilde{a}_j x_j &\leq a_j x_j + \Delta_j x_j \quad \forall x_j \geq 0 \\
\tilde{a}_j x_j &\leq a_j x_j - \Delta_j x_j \quad \forall x_j \leq 0
\end{aligned}$$

Note that both inequalities can be combined into a single inequality in terms of $|x_j|$.

$$\tilde{a}_j x_j \leq a_j x_j + \Delta_j |x_j| \quad \forall x_j$$

An absolute value formulation

As a result of the above worst-case analysis, solutions to the previous formulation of the original constraint with unknown-but-bounded coefficients $\tilde{a}_j$ can now be guaranteed by writing the following inequality without reference to $\tilde{a}_j$.

$$\sum_{j \in J} a_j x_j + \sum_{j \in J} \Delta_j |x_j| \leq b$$

A tolerance ...

In the above absolute value formulation it is usually too conservative to require that the original deterministic value of b cannot be loosened. Typically, a tolerance $\delta > 0$ is introduced to allow solutions x_j to violate the original right-hand side b by an amount of at most $\delta \max(1, |b|)$.

. . . relaxes the right-hand side

The term $\max(1, |b|)$ guarantees a positive increment of at least δ, even in case the right-hand side b is equal to zero. This modified right-hand side leads to the following *δ-tolerance* formulation where a solution x_j is feasible whenever it satisfies the following inequality.

$$\sum_{j \in J} a_j x_j + \sum_{j \in J} \Delta_j |x_j| \leq b + \delta \max(1, |b|)$$

The final formulation

This δ-tolerance formulation can be rewritten as a deterministic linear programming constraint by replacing the $|x_j|$ terms with nonnegative variables y_j, and requiring that $-y_j \leq x_j \leq y_j$. It is straightforward to verify that these last two inequalities imply that $y_j \geq |x_j|$. These two terms are likely to be equal when the underlying inequality becomes binding for optimal x_j values in a linear program. The final result is the following set of deterministic linear programming constraints, which captures the uncertainty reflected in the original constraint with unknown-but-bounded coefficients as presented at the beginning of this section.

$$\begin{aligned} \sum_{j \in J} a_j x_j + \sum_{j \in J} \Delta_j y_j &\leq b + \delta \max(1, |b|) \\ -y_j \leq x_j &\leq y_j \\ y_j &\geq 0 \end{aligned}$$

6.6 A probabilistic constraint

This section

This section considers the situation that occurs when the right-hand side of a linear constraint is a random variable. As will be shown, such a constraint can be rewritten as a purely deterministic constraint. Results pertaining to probabilistic constraints (also referred to as chance-constraints) were first published by Charnes and Cooper ([Ch59]).

Stochastic right-hand side

Consider the following linear constraint

$$\sum_{j \in J} a_j x_j \leq B$$

where $J = \{1, 2, \ldots, n\}$ and B is a random variable. A solution $x_j, j \in J$, is feasible when the constraint is satisfied for all possible values of B.

Acceptable values only

For open-ended distributions the right-hand side B can take on any value between $-\infty$ and $+\infty$, which means that there cannot be a feasible solution. If the distribution is not open-ended, suppose for instance that $B_{\min} \leq B \leq B_{\max}$, then the substitution of $B_{\min}$ for B results in a deterministic model. In most practical applications, it does not make sense for the above constraint to hold for all values of B.

A probabilistic constraint

Specifying that the constraint $\sum_{j \in J} a_j x_j \leq B$ must hold for all values of B is equivalent to stating that this constraint must hold with probability 1. In practical applications it is natural to allow for a small margin of failure. Such failure can be reflected by replacing the above constraint by an inequality of the form

$$\Pr\left[\sum_{j \in J} a_j x_j \leq B\right] \geq 1 - \alpha$$

which is called a *linear probabilistic constraint* or a *linear chance-constraint*. Here Pr denotes the phrase "Probability of", and α is a specified constant fraction ($\in [0, 1]$), typically denoting the maximum error that is allowed.

Deterministic equivalent

Consider the density function f_B and a particular value of α as displayed in Figure 6.2.

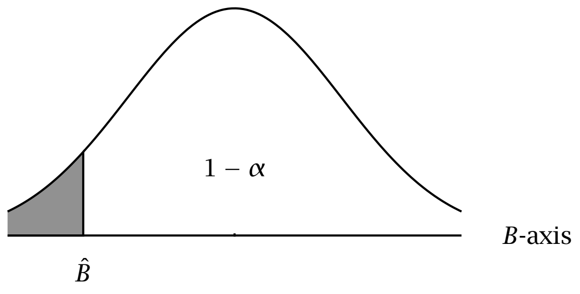


Figure 6.2: A density function f_B

A solution $x_j, j \in J$, is considered feasible for the above probabilistic constraint if and only if the term $\sum_{j \in J} a_j x_j$ takes a value beneath point $\hat{B}$. In this case a fraction $(1 - \alpha)$ or more of all values of B will be larger than the value of the term $\sum_{j \in J} a_j x_j$. For this reason $\hat{B}$ is called the *critical value*. The probabilistic constraint of the previous paragraph has therefore the following deterministic equivalent:

$$\sum_{j \in J} a_j x_j \leq \hat{B}$$

Computation of critical value

The critical value $\hat{B}$ can be determined by integrating the density function from $-\infty$ until a point where the area underneath the curve becomes equal to α. This point is then the value of $\hat{B}$. Note that the determination of $\hat{B}$ as described in this paragraph is equivalent to using the inverse cumulative distribution function of f_B evaluated at α. From probability theory, the cumulative distribution function F_B is defined by $F_B(x) = \Pr[B \le x]$. The value of F_B is the corresponding area underneath the curve (probability). Its inverse specifies for each particular level of probability, the point $\hat{B}$ for which the integral equals the probability level. The cumulative distribution function F_B and its inverse are illustrated in Figure 6.3.

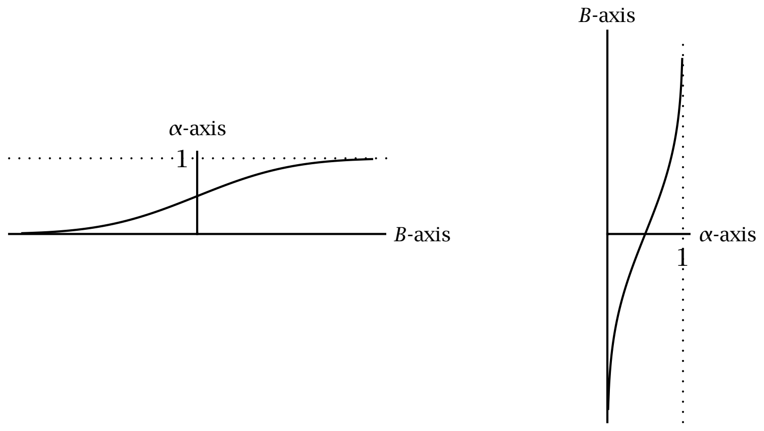


Figure 6.3: Cumulative distribution function F and its inverse.

Use AIMMS*-supplied function*

As the previous paragraph indicated, the critical $\hat{B}$ can be determined through the inverse of the cumulative distribution function. AIMMS supplies this function for a large number of distributions. For instance, when the underlying distribution is normal with mean 0 and standard deviation 1, then the value of $\hat{B}$ can be found as follows:

$$\hat{B} = \texttt{InverseCumulativeDistribution(Normal(0,1) ,}\alpha\texttt{)}$$

Example

Consider the constraint $\sum_j a_j x_j \le B$ with a stochastic right-hand side. Let $B = N(0,1)$ and $\alpha = 0.05$. Then the value of $\hat{B}$ based on the inverse cumulative distribution is -1.645. By requiring that $\sum_j a_j x_j \le -1.645$, you make sure that the solution x_j is feasible for 95% of all instances of the random variable B.

Overview of probabilistic constraints

The following figure presents a graphical overview of the four linear probabilistic constraints with stochastic right-hand sides, together with their deterministic equivalent. The shaded areas correspond to the feasible region of $\sum_{j \in J} a_j x_j$.

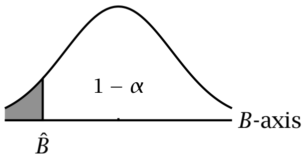
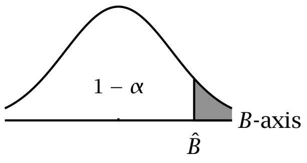
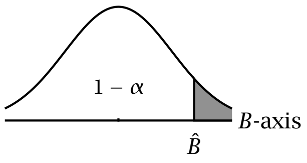
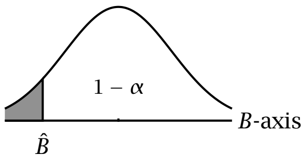
$\Pr\left[\sum_{j\in J} a_j x_j \le B\right] \ge 1-\alpha$ $\sum_{j\in J} a_j x_j \le \hat{B}$	
$\Pr\left[\sum_{j\in J} a_j x_j \le B\right] \le \alpha$ $\sum_{j\in J} a_j x_j \ge \hat{B}$	
$\Pr\left[\sum_{j\in J} a_j x_j \ge B\right] \ge 1-\alpha$ $\sum_{j\in J} a_j x_j \ge \hat{B}$	
$\Pr\left[\sum_{j\in J} a_j x_j \ge B\right] \le \alpha$ $\sum_{j\in J} a_j x_j \le \hat{B}$	

Table 6.1: Overview of linear probabilistic constraints

6.7 Summary

This chapter presented a number of techniques to transform some special models into conventional linear programming models. It was shown that some curve fitting procedures can be modeled, and solved, as linear programming models by reformulating the objective. A method to reformulate objectives which incorporate *absolute values* was given. In addition, a trick was shown to make it possible to incorporate a *minimax objective* in a linear programming model. For the case of a *fractional objective* with linear constraints, such as those that occur in financial applications, it was shown that these can be transformed into linear programming models. A method was demonstrated to specify a *range constraint* in a linear programming model. At the end of this chapter, it was shown how to reformulate constraints with a stochastic right-hand side to deterministic linear constraints.

Bibliography

[Ch59] A. Charnes and W.W. Cooper, *Change-constrained programming*, Management Science **6** (1959), 73–80.

[Ch62] A. Charnes and W.W. Cooper, *Programming with linear fractional functional*, Naval Research Logistics Quarterly **9** (1962), 181–186.

[Wi90] H.P. Williams, *Model building in mathematical programming*, 3rd ed., John Wiley & Sons, Chichester, 1990.