Oral exam questions in the course TMA521 Project Course Optimization

Torbjörn Larsson and Michael Patriksson

June 28, 2004

Contents

Part I

Lagrangian relaxation

Chapter 1

Global optimality conditions: zero duality gap

Consider the optimization problem to find

$$f^* := \underset{\boldsymbol{x}}{\text{infimum}}\ f(\boldsymbol{x}), \tag{1.1a}$$
$$\text{subject to } \boldsymbol{x} \in X, \tag{1.1b}$$
$$g_i(\boldsymbol{x}) \leq 0, \qquad i = 1, \dots, m, \tag{1.1c}$$

where $f : \mathbb{R}^n \mapsto \mathbb{R}$ and $g_i : \mathbb{R}^n \mapsto \mathbb{R}$ ($i = 1, 2, \dots, m$) are continuous functions, and $X \subseteq \mathbb{R}^n$ is closed.

For an arbitrary vector $\boldsymbol{\mu} \in \mathbb{R}^m$, we define the *Lagrange function*

$$L(\boldsymbol{x}, \boldsymbol{\mu}) := f(\boldsymbol{x}) + \sum_{i=1}^{m} \mu_i g_i(\boldsymbol{x}) = f(\boldsymbol{x}) + \boldsymbol{\mu}^{\mathrm{T}} \boldsymbol{g}(\boldsymbol{x}). \tag{1.2}$$

Let

$$q(\boldsymbol{\mu}) := \underset{\boldsymbol{x} \in X}{\text{infimum}}\ L(\boldsymbol{x}, \boldsymbol{\mu}) \tag{1.3}$$

be the *Lagrangian dual function*, defined by the infimum value of the Lagrange function over X when we have Lagrangian relaxed the explicit constraints with multiplier values $\boldsymbol{\mu}$; the *Lagrangian dual problem* is to

$$\underset{\boldsymbol{\mu}}{\text{maximize}}\ q(\boldsymbol{\mu}), \tag{1.4a}$$
$$\text{subject to } \boldsymbol{\mu} \geq \boldsymbol{0}^m. \tag{1.4b}$$

1.1 Global optimality condition imply primal optimality

Consider the following two-step procedure for solving the primal problem (1.1). We first solve the Lagrangian dual problem (1.4). Let $\boldsymbol{\mu}^*$ denote a dual opti-

mum. Thereafter, if possible, we generate an $\boldsymbol{x}^* \in \mathbb{R}^n$ which satisfies the *global optimality conditions*, that is,

$$\boldsymbol{\mu}^* \geq \boldsymbol{0}^m, \qquad \text{(Dual feasibility)} \tag{1.5a}$$
$$\boldsymbol{x}^* \in \arg\min_{\boldsymbol{x} \in X} L(\boldsymbol{x}, \boldsymbol{\mu}^*), \qquad \text{(Lagrangian optimality)} \tag{1.5b}$$
$$\boldsymbol{x}^* \in X, \quad \boldsymbol{g}(\boldsymbol{x}^*) \leq \boldsymbol{0}^m, \qquad \text{(Primal feasibility)} \tag{1.5c}$$
$$(\boldsymbol{\mu}^*)^{\mathrm{T}} \boldsymbol{g}(\boldsymbol{x}^*) = 0. \qquad \text{(Complementary slackness)} \tag{1.5d}$$

Show that such an $\boldsymbol{x}^*$ solves the primal problem. (Note that the first condition implies that $\boldsymbol{x}^* \in X$.)

1.2 Example

If the primal problem is convex, has an optimal solution, and also satisfies some constraint qualification, it is always possible to satisfy the conditions (1.5). Hence, the primal problem can always be solved by using the procedure described earlier.

Use the procedure to solve the linear (therefore convex) problem to

$$\begin{array}{lrl} \text{minimize} & z = & x_1 - 3x_2 \\ \text{subject to} & & -x_1 + 2x_2 \leq 6, \\ & & x_1 + x_2 \leq 5, \\ & & x_1, \quad x_2 \geq 0. \end{array}$$

Lagrangian relax the first constraint only.

1.3 A special case

Suppose that the Lagrangian relaxed problem in (1.3) has a *unique* minimum, and that we denote this vector $\boldsymbol{x}(\boldsymbol{\mu})$. How do the results above simplify? Provide sufficient conditions on the problem data [that is, on f, g_i ($i = 1, \dots, m$), and X] such that the Lagrangian relaxed problem in (1.3) has a unique minimum for every $\boldsymbol{\mu} \geq \boldsymbol{0}^m$. This then holds in particular for an (a priori unknown) optimal solution $\boldsymbol{\mu}^*$.

[*Note:* If $\boldsymbol{\mu}^* \geq \boldsymbol{0}^m$ and $\boldsymbol{g}(\boldsymbol{x}^*) \leq \boldsymbol{0}^m$ then $(\boldsymbol{\mu}^*)^{\mathrm{T}} \boldsymbol{g}(\boldsymbol{x}^*) = 0$ if and only if $\mu_i^* g_i(\boldsymbol{x}^*) = 0$ for $i = 1, \dots, m$. The complementarity conditions (1.5d) in the global optimality conditions can therefore also be expressed as

$$\mu_i^* g_i(\boldsymbol{x}^*) = 0, \qquad i = 1, \dots, m.$$

It is usually of advantage to utilize this form when searching for a vector $\boldsymbol{x}^* \in \mathbb{R}^n$ satisfying the global optimality conditions given an optimal dual solution $\boldsymbol{\mu}^*$. The reason is that the separate complementarity conditions $\mu_i^* g_i(\boldsymbol{x}^*) = 0$,

$i = 1, \ldots, m$, directly provide information about which constraints $g_i(\boldsymbol{x}) \leq 0$, $i = 1, \ldots, m$, which an $\boldsymbol{x}^*$ sought must fulfill with equality; the aggregated complementarity condition $(\boldsymbol{\mu}^*)^{\mathrm{T}} \boldsymbol{g}(\boldsymbol{x}^*) = 0$ alone does not provide this information.

We also note that if the original problem has the form

$$f^* := \underset{\boldsymbol{x}}{\operatorname{infimum}}\ f(\boldsymbol{x}), \tag{1.6a}$$

$$\text{subject to } \boldsymbol{x} \in X, \tag{1.6b}$$

$$h_j(\boldsymbol{x}) = 0, \qquad j = 1, \ldots, \ell, \tag{1.6c}$$

where $h_j : \mathbb{R}^n \mapsto \mathbb{R}$ $(j = 1, \ldots, \ell)$ are continuous, then the global optimality conditions will reduce to the following system in terms of the multipliers λ_j $(j = 1, \ldots, \ell)$ and the associated Lagrange function $L(\boldsymbol{x}, \boldsymbol{\lambda}) := f(\boldsymbol{x}) + \sum_{j=1}^{\ell} \lambda_j h_j(\boldsymbol{x}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{h}(\boldsymbol{x})$:

$$\boldsymbol{x}^* \in \arg\min_{\boldsymbol{x} \in X} L(\boldsymbol{x}, \boldsymbol{\lambda}^*), \qquad \text{(Lagrangian optimality)} \tag{1.7a}$$

$$\boldsymbol{x}^* \in X, \quad \boldsymbol{h}(\boldsymbol{x}^*) = \boldsymbol{0}^{\ell}, \qquad \text{(Primal feasibility)} \tag{1.7b}$$

since the complementarity conditions then always are satisfied.]

Chapter 2

Additional properties of the Lagrangian dual problem

Consider the problem (1.1) with the properties stated together with its definition.

2.1 The appearance of the Lagrangian dual function

Suppose now that X is finite (for example, X consists of a finite number of integer vectors). Denote the elements in X by $\boldsymbol{x}^p$, $p = 1, \ldots, P$. Show that the dual objective function q is piece-wise linear. How many pieces does it have, at most? Why is it not always built up by this many pieces?

[*Note:* This result holds regardless of the properties of f and $\boldsymbol{g}$.]

2.2 Example

Illustrate the above result for the linear 0/1 problem to find

$$\begin{array}{rll} z^* = & \text{maximum} & z = 5x_1 + 8x_2 + 7x_3 + 9x_4 \\ & \text{subject to} & \quad 3x_1 + 2x_2 + 2x_3 + 4x_4 \leq 5 \\ & & \quad 2x_1 + x_2 + 2x_3 + x_4 = 3 \\ & & \quad x_1\ ,\ x_2\ ,\ x_3\ ,\ x_4 \in \{0, 1\}, \end{array}$$

where the first constraint is considered complicating and is to be Lagrangian relaxed.

2.3 Linear programming

Suppose that f and all functions g_i are affine, and that X is a polytope (that is, a bounded polyhedron). Show that the dual objective function is piece-wise linear also in this case. What is the maximum number of segments of the Lagrangian dual function?

Chapter 3

Lagrangian relaxation and convexification

Consider again the framework of Section 1. There, we concluded that for convex problems a primal optimal solution $\boldsymbol{x}^*$ can be generated through a procedure in which first an optimal dual solution $\boldsymbol{\mu}^*$ is found, after which $\boldsymbol{x}^*$ is generated as a solution to the global optimality conditions (1.5).

3.1 Non-convex example

For non-convex problems it is not guaranteed that this procedure can be used to derive a primal optimal solution. Illustrate this fact by considering the problem to

$$\begin{aligned} \text{minimize } & f(x) = -2x_1 + x_2 \\ \text{subject to } & x_1 + x_2 \le 3, \\ & \boldsymbol{x} \in X = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 4 \\ 4 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}. \end{aligned}$$

Which of the global optimality conditions in (1.5) cannot be fulfilled? Calculate the size of the duality gap, that is, $\Gamma = f^* - q^*$.

3.2 Convexified version

Consider the problem to

$$\begin{aligned} \text{minimize } & f(x) = -2x_1 + x_2 \\ \text{subject to } & x_1 + x_2 \le 3, \\ & \boldsymbol{x} \in X^c = \{ \boldsymbol{x} \in \mathbb{R}^2 \mid 0 \le x_j \le 4;\ j = 1, 2 \}, \end{aligned}$$

which is a *convexified* version of the previous problem. Solve this problem and evaluate its optimal objective value, f_c^*. What is the relation between f_c^* and the optimal dual value q^*?

[*Note:* This relation can be shown to be valid generally.]

Chapter 4

Global optimality conditions: non-zero duality gap

Consider the optimization problem (1.1), the Lagrange function (1.2), and the associated Lagrangian dual problem (1.3), under the conditions stated in the beginning of Section 1.

Let $(\boldsymbol{x}^*, \boldsymbol{\mu}^*)$ be a pair of primal–dual optimal solutions to the respective primal and dual problems. With f^* and q^* being their respective values we define the *duality gap* as $\Gamma := f^* - q^*$, which is non-negative by weak duality.

4.1 Global optimality condition imply primal optimality

We extend the procedure outlined in Section 1.1, as follows. We first solve the Lagrangian dual problem (1.4). Let $\boldsymbol{\mu}^*$ denote a dual optimum. Thereafter, if possible, we generate an $\boldsymbol{x}^* \in \mathbb{R}^n$ which satisfies the *global optimality conditions*, that is,

$$\begin{aligned} f(\boldsymbol{x}) + \boldsymbol{\mu}^{\mathrm{T}} \boldsymbol{g}(\boldsymbol{x}) &\le q(\boldsymbol{\mu}) + \varepsilon && \text{(Lagrangian } \varepsilon\text{-optimality)} && (4.1a) \\ \boldsymbol{g}(\boldsymbol{x}) &\le \mathbf{0}^m && \text{(Primal feasibility)} && (4.1b) \\ \boldsymbol{\mu}^{\mathrm{T}} \boldsymbol{g}(\boldsymbol{x}) &\ge -\delta && \text{(Lagrangian } \delta\text{-complementarity)} && (4.1c) \\ \varepsilon + \delta &\le \Gamma && \text{(Perturbation condition, I)} && (4.1d) \\ \varepsilon, \delta &\ge 0 && \text{(Perturbation condition, II)} && (4.1e) \end{aligned}$$

Establish the equivalence between the following two statements:

1. there exists a pair (ε, δ) of perturbations such that the pair $(\boldsymbol{x}^*, \boldsymbol{\mu}^*)$ satisfies (4.1);
2. the pair $(\boldsymbol{x}^*, \boldsymbol{\mu}^*)$ defines a primal–dual optimal solution.

4.2 An disaggregated version of the global optimality conditions

Consider the following extension of the system (4.1):

$$\begin{aligned} f(\boldsymbol{x}) + \boldsymbol{\mu}^{\mathrm{T}} \boldsymbol{g}(\boldsymbol{x}) &\le q(\boldsymbol{\mu}) + \varepsilon && \text{(Lagrangian } \varepsilon\text{-optimality)} && (4.2a) \\ \boldsymbol{g}(\boldsymbol{x}) &\le \mathbf{0}^m && \text{(Primal feasibility)} && (4.2b) \\ \mu_i g_i(\boldsymbol{x}) &\ge -\delta_i,\ i = 1, \dots, m && \text{(Lagrangian } \delta\text{-complementarity)} && (4.2c) \\ \varepsilon + \sum_{i=1}^m \delta_i &\le \Gamma && \text{(Perturbation condition, I)} && (4.2d) \\ \varepsilon \ge 0,\ \boldsymbol{\delta} &\ge \mathbf{0}^m && \text{(Perturbation condition, II)} && (4.2e) \end{aligned}$$

Show that this version of the global optimality conditions is equivalent to the original one, in that we can also characterize a primal–dual optimal pair with a pair $(\varepsilon, \boldsymbol{\delta})$ in the system (4.2). ***Hint:*** It is the easiest to establish the equivalence directly between the two systems.

[*Note:* The appearance of the system (4.2) is that of a disaggregated version of the system (4.1), wherein we measure the complementarity fulfillment for each constraint individually rather than lumped together into a sum. Utilizing this disaggregation can some times be of advantage when devising a Lagrangian heuristic for the problem (1.1), especially when it is of the column generation variety.]

4.3 Example

Solve the problem to

$$\begin{aligned} \text{minimize } & f(x) = -2x_1 + x_2 \\ \text{subject to } & x_1 + x_2 \le 3, \\ & \boldsymbol{x} \in X = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 4 \\ 4 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\} \end{aligned}$$

by first solving the Lagrangian dual problem associated with the Lagrangian relaxation of the linear constraint, and then generate the set of optimal primal solutions through the use of the system (4.1). For which values of ε and δ can we derive a primal–dual optimal solution?

Chapter 5

The network design problem

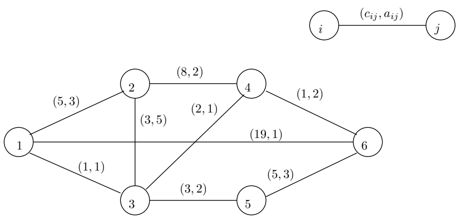
5.1 The minimum spanning tree problem

Formulate the minimum spanning tree problem (MST) as a network flow problem. [*Hint*: consider node 1 as a sink and all other nodes as sources with strength 1.]

Formulate also the MST problem as a binary, integer programming problem.

5.2 Numerical example

Consider the graph below.



Provide *all* the spanning trees of this graph explicitly. Calculate the sum of c_{ij} and a_{ij} for each tree. Which ones are feasible with respect to the *budget*

constraint

$$\sum_{(i,j)\in\mathcal{T}} a_{ij} \leq 10$$

(where $\mathcal{T}$ denotes a collection of links forming a spanning tree)? Which ones are optimal (minimal) with respect to the link costs c_{ij}?

The MST problem under a budget constraint is often referred to as the *network design problem* (NDP).

5.3 Primal heuristics

Provide a *local search* heuristic for the NDP which improves a feasible solution.

Apply it to the above example.

5.4 Lagrangian heuristics

Provide a *Lagrangian relaxation* based algorithm for the NDP. In particular, choose a suitable Lagrangian relaxation, a description of how to solve each Lagrangian subproblem, a primal feasibility heuristic, and a complete Lagrangian scheme including a suitable step length formula.

Apply it to the above example.

Chapter 6

A Lagrangian heuristic

Consider the linear 0/1 problem to find

$$\begin{array}{rrcrcrcrcll}
z^* = \text{maximum } z = & 5x_1 & + & 8x_2 & + & 7x_3 & + & 9x_4 & & & \\
\text{subject to} & 3x_1 & + & 2x_2 & + & 3x_3 & + & 3x_4 & \leq & 6 & (1)\\
 & 2x_1 & + & 3x_2 & + & 3x_3 & + & 4x_4 & \leq & 5 & (2)\\
 & x_1 & & & + & x_3 & & & = & 1 & \\
 & & & x_2 & & & + & x_4 & = & 1 & \\
 & x_1 & , & x_2 & , & x_3 & , & x_4 & \in & \{0,1\}. &
\end{array}$$

6.1 Lagrangian relaxation

Lagrangian relax the constraints (1) and (2) with multipliers $\mu_1, \mu_2 \geq 0$. Formulate the relaxed problem and the corresponding Lagrangian dual problem. Analyze its properties.

6.2 Weak duality

Evaluate the Lagrangian dual objective function at $\boldsymbol{\mu} = (1,1)^{\mathrm{T}}$. Relate this value to z^*.

6.3 Subgradients

Calculate a subgradient to the Lagrangian dual objective function at $\boldsymbol{\mu} = (1,1)^{\mathrm{T}}$.

6.4 Primal feasibility heuristic

Find a feasible solution to the original problem by suitably modifying the solution to the Lagrangian subproblem at $\boldsymbol{\mu} = (1,1)^{\mathrm{T}}$. Relate this value to z^*.

6.5 Subgradient optimization

How are the multipliers in the subgradient optimization method updated? How can the results in the previous assignment aid in the calculation of the step length.

6.6 Strong duality

What is the relation between the value of the Lagrangian dual function and the value of z^*? Is it likely that a positive duality gap is present? Why/why not?

6.7 Continuous relaxation

What is the relationship between the optimal value of the Lagrangian dual problem and that of the continuous (LP) relaxation of the original problem? Motivate!

Chapter 7

Subgradients and optimality for unconstrained Lagrangian duals

Consider the primal problem

$$f^* := \operatorname*{infimum}_{\boldsymbol{x}} f(\boldsymbol{x}), \tag{7.1a}$$
$$\text{subject to } \boldsymbol{x} \in X, \tag{7.1b}$$
$$h_j(\boldsymbol{x}) = 0, \qquad j = 1, \dots, \ell, \tag{7.1c}$$

where $f : \mathbb{R}^n \mapsto \mathbb{R}$ and $h_j : \mathbb{R}^n \mapsto \mathbb{R}$ $(j = 1, 2, \dots, \ell)$ are continuous functions, and $X \subseteq \mathbb{R}^n$ is closed.

For an arbitrary vector $\boldsymbol{\lambda} \in \mathbb{R}^\ell$, we define the *Lagrange function*

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) := f(\boldsymbol{x}) + \sum_{j=1}^{\ell} \lambda_j h_j(\boldsymbol{x}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{h}(\boldsymbol{x}). \tag{7.2}$$

Let

$$q(\boldsymbol{\lambda}) := \operatorname*{infimum}_{\boldsymbol{x} \in X} L(\boldsymbol{x}, \boldsymbol{\lambda}) \tag{7.3}$$

be the *Lagrangian dual function*, defined by the infimum value of the Lagrange function over X when we have Lagrangian relaxed the explicit constraints with multiplier values $\boldsymbol{\lambda}$; the *Lagrangian dual problem* is to

$$\operatorname*{maximize}_{\boldsymbol{\lambda}} q(\boldsymbol{\lambda}), \tag{7.4a}$$
$$\text{subject to } \boldsymbol{\lambda} \in \mathbb{R}^\ell. \tag{7.4b}$$

We suppose that X is non-empty, closed and bounded so that the "infimum" can be replaced by "minimum" in (7.1), (7.3), and both the primal and dual

objective functions in (7.1) and (7.4) attain their optimal values. Moreover, then q is finite, continuous, and concave on the whole set $\mathbb{R}^\ell$, so that (7.4) is a "well-behaved" convex problem.

7.1 Subgradients of concave functions, I

Let $\bar{\boldsymbol{\lambda}} \in \mathbb{R}^\ell$. Define a subgradient, $\bar{\boldsymbol{\gamma}}$, to q at $\bar{\boldsymbol{\lambda}}$. Also define the entire subdifferential, $\partial q(\bar{\boldsymbol{\lambda}})$ to q at $\bar{\boldsymbol{\lambda}}$.

7.2 Subgradients of concave functions, II

Let $\boldsymbol{x}(\bar{\boldsymbol{\lambda}})$ be an optimal solution to the Lagrangian subproblem defining the value of $q(\bar{\boldsymbol{\lambda}})$. Show that the vector $\boldsymbol{h}(\boldsymbol{x}(\bar{\boldsymbol{\lambda}})) \in \partial q(\bar{\boldsymbol{\lambda}})$; that is, show that the vector $\boldsymbol{h}(\boldsymbol{x}(\bar{\boldsymbol{\lambda}}))$ of constraint function values at the subproblem solution $\boldsymbol{x}(\bar{\boldsymbol{\lambda}})$ is a subgradient of q at $\bar{\boldsymbol{\lambda}}$.

7.3 Optimality

Show that if, for some $\boldsymbol{\lambda}^* \in \mathbb{R}^\ell$, it holds that $0^\ell \in \partial q(\boldsymbol{\lambda}^*)$ then $\boldsymbol{\lambda}^*$ is globally optimal in the Lagrangian dual problem (7.4).

7.4 Optimality, numerical example

Illustrate the above result on the LP problem to

$$\begin{array}{llcrcrcl}
\text{minimize } z = & x_1 & - & 3x_2 & & \\
\text{subject to} & -x_1 & + & 2x_2 & = & 6 \\
 & x_1 & + & x_2 & \leq & 5 \\
 & x_1 & , & x_2 & \geq & 0,
\end{array}$$

where the first constraint is considered complicating and is therefore Lagrangian relaxed. At $\lambda^* = 4/3$ there are alternative Lagrangian subproblem optimal solutions, which yield alternative subgradients of the Lagrangian dual function. There are, especially, exactly *two* extreme solutions to the Lagrangian subproblem, which both yield extreme subgradients (that is, subgradients of q at λ^* that are not non-trivial convex combinations of other subgradients of q at λ^*). Find these two extreme subgradients. Characterize the subdifferential $\partial q(\lambda^*)$. Verify that λ^* is globally optimal in the Lagrangian dual problem.

Chapter 8

Subgradients and optimality for constrained Lagrangian duals

Consider the inequality constrained problem (1), the Lagrange function (1.2), and Lagrangian dual problem (1.4), where we assume that the functions f and g_i $(i = 1, \dots, m)$ are continuous and X non-empty, closed and bounded. It follows that the Lagrangian dual function q is finite, continuous, and concave on $\mathbb{R}^m$, and that (1.4) therefore is a "well-behaved" convex problem. Further, both the primal and dual problems have non-empty, closed and bounded optimal solution sets.

8.1 Optimality

Let $\boldsymbol{\mu}^* \geq \mathbf{0}^m$. Show that if there exists a subgradient $\boldsymbol{\gamma}^*$ to q at $\boldsymbol{\mu}^*$ that satisfies

$$\boldsymbol{\gamma}^* \leq \mathbf{0}^m, \qquad \text{and} \qquad (\boldsymbol{\mu}^*)^{\mathrm{T}} \boldsymbol{\gamma}^* = 0$$

(we then say that the vectors $\boldsymbol{\mu}^*$ and $\boldsymbol{\gamma}^*$ are *complementary* to each other) then $\boldsymbol{\mu}^*$ is an optimal solution to the Lagrangian dual problem (1.4).

8.2 Numerical example

Consider the LP problem to

$$\begin{array}{llcrcrl}
\text{minimize } z = & x_1 & - & 3x_2 & & & \\
\text{subject to} & -x_1 & + & 2x_2 & \leq & 6 & \quad (1) \\
 & -x_1 & - & 2x_2 & \leq & -7 & \quad (2) \\
 & x_1 & + & x_2 & \leq & 5 & \\
 & x_1 & , & x_2 & \geq & 0, &
\end{array}$$

where the constraints (1) and (2) are considered complicating and therefore are Lagrangian relaxed. At $\boldsymbol{\mu}^* = (4/3, 0)^{\mathrm{T}}$ there exist exactly *two* extreme subgradients of q; find them. Characterize the subdifferential $\partial q(\boldsymbol{\mu}^*)$. Find a subgradient, $\boldsymbol{\gamma}^* \in \partial q(\boldsymbol{\mu}^*)$, to q at $\boldsymbol{\mu}^*$ which verifies that $\boldsymbol{\mu}^* = (4/3, 0)^{\mathrm{T}}$ indeed is Lagrangian dual optimal.

Graphically illustrate in $\mathbb{R}^2$ the relations between the vectors $\boldsymbol{\mu}^*$, $\boldsymbol{\gamma}^*$, and the two extreme subgradients.

Chapter 9

Lagrangian relaxation and convexification

Consider the primal problem to

$$\underset{\boldsymbol{x}}{\text{minimize}} \ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}, \tag{9.1a}$$
$$\text{subject to } \boldsymbol{x} \in X, \tag{9.1b}$$
$$\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}, \tag{9.1c}$$

where $X \subseteq \mathbb{R}^n_+$, and the linear constraints are considered complicating. Consider the Lagrangian dual problem to find

$$q^* = \underset{\boldsymbol{\mu} \geq \boldsymbol{0}^m}{\text{maximum}} \ q(\boldsymbol{\mu}), \tag{9.2}$$

where

$$q(\boldsymbol{\mu}) = \underset{\boldsymbol{x} \in X}{\text{minimum}} \ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}). \tag{9.3}$$

Suppose that the set X is finite. (For example, it may contain a finite number of integer vectors in $\mathbb{R}^n$.) Denote the vectors in X by $\boldsymbol{x}^i$, $i = 1, \ldots, P$.

9.1 Lagrangian dual problem, I

Show that the Lagrangian dual problem can equivalently be formulated as the problem to find

$$\begin{aligned} q^* = \text{maximum} \ & z, \\ \text{subject to} \ & z \leq (\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^i)^{\mathrm{T}}\boldsymbol{\mu} + \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^i, \qquad i = 1, \ldots, P, \\ & \boldsymbol{u} \geq \boldsymbol{0}^m. \end{aligned}$$

9.2 Lagrangian dual problem, II

By using the previous formulation, show that the formulation also can be written as follows:

$$\begin{aligned} q^* = \text{minimum} \ & \sum_{i=1}^{P} (\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^i)\mu_i, \\ \text{subject to} \ & \sum_{i=1}^{P} (\boldsymbol{A}\boldsymbol{x}^i)\lambda_i \geq \boldsymbol{b}, \\ & \sum_{i=1}^{P} \mu_i = 1, \\ & \boldsymbol{\mu} \geq \boldsymbol{0}^m. \end{aligned}$$

9.3 Lagrangian dual problem, III

By using the previous formulation, show that the formulation also can be written as follows:

$$\begin{aligned} q^* = \underset{\boldsymbol{x}}{\text{minimum}} \ & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}, \\ \text{subject to} \ & \boldsymbol{x} \in \operatorname{conv} X, \\ & \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}. \end{aligned}$$

9.4 Strength of Lagrangian relaxation, I

Which relation holds between X and $\operatorname{conv} X$? Which relation holds between z^* and q^*?

9.5 Numerical example

Consider the linear 0/1 problem to find

$$\begin{array}{rlrcrcrcrcl} z^* = & \text{maximum } z = & 5x_1 & + & 8x_2 & + & 7x_3 & + & 9x_4 & & \\ & \text{subject to} & 3x_1 & + & 2x_2 & + & 2x_3 & + & 4x_4 & \leq & 5 \\ & & 2x_1 & + & x_2 & + & 2x_3 & + & x_4 & = & 3 \\ & & x_1 & , & x_2 & , & x_3 & , & x_4 & \in & \{0,1\}, \end{array}$$

where the first constraint is considered complicating and is Lagrangian relaxed. Utilized the result in Section 9.2 to find the Lagrangian dual optimal value.

9.6 Strength of Lagrangian relaxation, II

Make the alternative assumption that the set X is a polytope (a bounded polyhedron), with extreme points $\boldsymbol{x}^i$, $i = 1, \ldots, P$. (The original problem hence is a linear program.) Motivate why the above reformulations are valid also under this assumption. What relation holds between X and $\operatorname{conv} X$? Which relation holds between z^* and q^*? What then is the problem in Section 9.2?

[*Note:* The conclusions in Section 9.4 is fundamental for the use of Lagrangian relaxation in integer programming. It is usually summarized as follows: Lagrangian relaxing a group of constraints is equivalent to convexifying the set defined by the constraints that are *not* Lagrangian relaxed. The problems in Sections 9.1 and 9.2 can normally not be formulated and solved, since the number P is usually extremely large. The problem in Section 9.2 can however be attacked using *column generation*.]

Chapter 10

Lagrangian relaxation vs. continuous relaxation

Suppose that the linear integer program

$$\begin{aligned} z^* = \text{minimum} \quad & z = \mathbf{c}^{\mathrm{T}}\mathbf{x}, \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = b, \\ & \mathbf{D}\mathbf{x} \leq \mathbf{e} \\ & \mathbf{x} \geq \mathbf{0}^n \\ & \mathbf{x} \text{ integer} \end{aligned}$$

is attacked by means of Lagrangian relaxation of the linear equality constraints. The relaxed problem's optimal value is then given by

$$\begin{aligned} q(\boldsymbol{\lambda}) = \text{minimum} \quad & \mathbf{c}^{\mathrm{T}}\mathbf{x} + \boldsymbol{\lambda}^{\mathrm{T}}(\mathbf{A}\mathbf{x} - \mathbf{b}) \\ \text{subject to} \quad & \mathbf{D}\mathbf{x} \leq \mathbf{e} \\ & \mathbf{x} \geq \mathbf{0}^n \\ & \mathbf{x} \text{ integer,} \end{aligned}$$

and the Lagrangian dual problem is to find

$$q^* = \underset{\boldsymbol{\lambda}}{\text{maximum}}\ q(\boldsymbol{\lambda}).$$

10.1 Integrality property

What does it mean that the Lagrangian relaxed problem has the *integrality property*? Give examples of Lagrangian relaxed problems that have (respectively, do not have) the integrality property. What methods are normally used to solve Lagrangian relaxed problems that have (respectively, do not have) the integrality property? Which one of the two types of problems normally is the most difficult to solve?

10.2 Bounds on the optimal value

Consider the continuous relaxation of the original integer program:

$$\begin{aligned} z^*_{LP} = \text{minimum} \quad & z = \mathbf{c}^{\mathrm{T}}\mathbf{x}, \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, \\ & \mathbf{D}\mathbf{x} \leq \mathbf{e}, \\ & \mathbf{x} \geq \mathbf{0}^n, \end{aligned}$$

and let $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ denote a dual optimum.

Show that the lower bound on z^* that is given by the Lagrangian dual function is always at least as good as that produced by the continuous relaxation. (In other words, show that $z^* \geq q^* \geq z^*_{LP}$.)

10.3 Comparative strength of relaxation

Show that if the Lagrangian relaxed problem has the integrality property then the two bounds are equal. Further, show that the partial LP dual solution $\boldsymbol{\lambda}^*$ is a Lagrangian dual optimal solution. Motivate why the Lagrangian dual problem can be expected to give stronger bounds than the continuous relaxation whenever the relaxed problem does not have the integrality property.

[*Note*: By the above result follows that Lagrangian relaxation, together with an algorithm for the (approximate) solution of the Lagrangian dual problem, can be used to find an (approximate) optimal solution to the LP relaxation of a linear integer program. Especially for some 0/1 problems this is a more effective means to attack the problem than to use traditional LP techniques. The reason is at least two-fold: (1) Methods that are based on Lagrangian relaxation can often utilize problem structure better than LP techniques. (This is the case for network-type problems, where a Lagrangian relaxation method typically only relies on the original data, which also can be very efficiently stored and manipulated.) (2) Continuous relaxations of 0/1 problems are often (heavily) degenerated, which hampers the efficiency of the simplex method. (As a side note we also remark that it has been observed empirically that it is often best to use the *dual* simplex method for solving the LP relaxation of a 0/1 problem, because it tends to fair better under degeneracy than the primal simplex method.)]

Chapter 11

Construction of strong LP relaxations

The Representation Theorem can some times be used to construct alternative formulations of discrete optimization problems.

Consider the problem to

$$\text{(P)} \qquad \begin{aligned} \text{minimize} \quad & \mathbf{c}^{\mathrm{T}}\mathbf{x}, \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b}, \\ & \mathbf{x} \in X \subseteq \{0,1\}^n, \end{aligned}$$

and suppose the set X contains P integer vectors, say $\mathbf{x}^i$, $i = 1, \dots, P$. By utilizing that

$$X = \left\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \sum_{i=1}^{P} \lambda_i \mathbf{x}^i; \quad \sum_{i=1}^{P} \lambda_i = 1; \quad \lambda_i \in \{0,1\},\ i = 1, \dots, P \right\}$$

we can write the problem equivalently as that to

$$\text{(IMP)} \qquad \begin{aligned} \text{minimize} \quad & \sum_{i=1}^{P} (\mathbf{c}^{\mathrm{T}}\mathbf{x}^i)\, \lambda_i, \\ \text{subject to} \quad & \sum_{i=1}^{P} (\mathbf{A}\mathbf{x}^i)\, \lambda_i \geq \mathbf{b}, \\ & \sum_{i=1}^{P} \lambda_i = 1, \\ & \lambda_i \in \{0,1\}, \quad i = 1, \dots, P. \end{aligned}$$

11.1 Numerical example

Consider the problem to

$$\begin{array}{lrcrcrcrcll} \text{minimize} & 3x_1 & + & 2x_2 & + & 4x_3 & + & 2x_4 & & & \\ \text{subject to} & 2x_1 & + & 3x_2 & + & 4x_3 & + & 5x_4 & \geq & 7 & (1) \\ & x_1 & + & x_2 & + & x_3 & + & 2x_4 & \leq & 3 & (2) \\ & x_1 & + & x_2 & & & & & = & 1 & (3) \\ & & & & & x_3 & + & x_4 & = & 1 & (4) \\ & x_1 & , & x_2 & , & x_3 & , & x_4 & \in & \{0/1\}. & (5) \end{array}$$

Let the constraints (1), (2), and (5) define the ground set on which the Representation Theorem is used (that is, the set X above). Formulate the equivalent problem. Solve it by inspection. (***Hint:*** The set X contains five integer vectors.)

11.2 Column generation

Consider again the original problem (P). Describe how the continuous relaxation of the problem (IMP), that is, the problem to

$$\text{(MP)} \qquad \begin{aligned} \text{minimize} \quad & \sum_{i=1}^{P} (\mathbf{c}^{\mathrm{T}}\mathbf{x}^i)\, \lambda_i, \\ \text{subject to} \quad & \sum_{i=1}^{P} (\mathbf{A}\mathbf{x}^i)\, \lambda_i \geq \mathbf{b}, \\ & \sum_{i=1}^{P} \lambda_i = 1, \\ & \lambda_i \geq 0, \quad i = 1, \dots, P, \end{aligned}$$

can be solved by the use of column generation. Show generally that the column generation problem is a Lagrangian relaxation of the problem (P) where the linear constraints are Lagrangian relaxed.

11.3 Convexification

To solve the continuous relaxation (MP) is obviously equivalent to solving the convexified problem to

$$\text{(CP)} \qquad \begin{aligned} \text{minimize} \quad & \mathbf{c}^{\mathrm{T}}\mathbf{x}, \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b}, \\ & \mathbf{x} \in \operatorname{conv} X, \end{aligned}$$

where

$$\operatorname{conv} X = \left\{ \mathbf{x} \mid \mathbf{x} = \sum_{i=1}^{P} \lambda_i \mathbf{x}^i; \quad \sum_{i=1}^{P} \lambda_i = 1; \quad \lambda_i \geq 0,\ i = 1, \dots, P \right\}.$$

If the set X has the description $X = \{\, \boldsymbol{x} \in \{0,1\}^n \mid \boldsymbol{Bx} \geq \boldsymbol{d} \,\}$ then the original problem has the continuous relaxation to

$$
\begin{array}{lll}
 & \text{minimize} & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}, \\
 & \text{subject to} & \boldsymbol{Ax} \geq \boldsymbol{b}, \\
\text{(LP)} & & \boldsymbol{Bx} \geq \boldsymbol{d}, \\
 & & 1 \geq x_j \geq 0, \quad j = 1, \ldots, n.
\end{array}
$$

When do the two continuous relaxations (LP) and (MP) have the same optimal objective value, and when do they normally differ? In the latter case, which one of the two relaxations is the strongest? In general, then, which one of the two formulations of the problem, (P) or (IMP), is in this sense the strongest formulation?

[*Hint:* Construct the continuous relaxations (LP) and (CP) for the numerical example

$$
\begin{array}{lrcrcl}
\text{minimize} & 2x_1 & + & x_2 & & \\
\text{subject to} & -x_1 & + & x_2 & \geq & 0 \\
 & 2x_1 & + & 2x_2 & \geq & 1 \\
 & x_1 & , & x_2 & \in & \{0,1\}
\end{array}
$$

with $X = \{\, (x_1, x_2) \in \{0,1\}^2 \mid 2x_1 + 2x_2 \geq 1 \,\}$.]

[*Note*: If (MP) has a higher optimal value than (LP) we then say that (MP) is a *strong linear programming formulation* of (P). Strong LP formulations can be obtained by other means as well. A common way is to extend a given formulation with constraints that are redundant in the integer program but which are *not* redundant in the continuous relaxation. A classic example is the *Gomory cut*.]

11.4 Pros and cons

Which are the pros and cons of solving (MP) [which provides an optimal solution to (CP)] and (LP)?

Chapter 12

A Lagrangian heuristic for the capacitated localization problem

Consider the *capacitated localization problem*, formulated as a mixed integer problem:

$$
\begin{array}{llcll}
\text{minimize} & \displaystyle\sum_{i=1}^{m} f_i y_i & + & \displaystyle\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij} x_{ij} & , \\
\text{subject to} & \displaystyle\sum_{j=1}^{n} x_{ij} & \leq & S_i y_i, & i = 1, \ldots, m, \\
 & \displaystyle\sum_{i=1}^{m} x_{ij} & = & D_j, & j = 1, \ldots, n, \qquad (*) \\
 & y_i & = & 0/1, & i = 1, \ldots, m, \\
 & x_{ij} & \geq & 0, & i = 1, \ldots, m, \quad j = 1, \ldots, n,
\end{array}
$$

where the variables y_i, $i = 1, \ldots, m$, describes logical decisions concerning the location of sources, and x_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$, describe quantities of items transported.

12.1 A redundant constraint

Show that the constraint

$$\sum_{i=1}^{m} S_i y_i \geq \sum_{j=1}^{n} D_j$$

is a consequence of the original constraints (that is, every feasible solution satisfies this constraint), whence it can be added to the model without affecting its solution.

12.2 Lagrangian relaxation

Suppose we add this redundant constraint to the original model and solve the resulting problem heuristically by Lagrangian relaxing the constraints $(*)$ and utilized subgradient optimization on the Lagrangian dual problem. Describe the solution process! Which are the main properties of the dual problem? Why? Can there be a positive duality gap?

12.3 Feasibility heuristics

For this problem it is easy to construct feasible solutions based on the optimal solution to each Lagrangian subproblem. Explain how!

12.4 Lower bounds

Compare the strength of the lower bound obtained from the Lagrangian dual formulation compared with that of the continuous (LP) relaxation of the problem. How is the strength of the lower bound affected if we do *not* add the above redundant constraint? Motivate!

[*Note*: Note that the added constraint is redundant in the original model, but that it is not redundant in the Lagrangian relaxed problem. We thereby illustrate the fact that it is often both possible and advantageous to strengthen a relaxation by means of adding constraints that are redundant in the original formulation.]

Chapter 13

Everett's Theorem

For the problem

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}^m \quad (P) \\ & \boldsymbol{x} \in X, \end{array}$$

where $X \subseteq \mathbb{R}^n$ is closed, and $f : \mathbb{R}^n \to \mathbb{R}$ and $\boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^m$ are continuous, it holds that if $\boldsymbol{x}^* \in \mathbb{R}^n$ and $\boldsymbol{\mu}^* \in \mathbb{R}^m_+$ satisfy the *global optimality conditions*, that is, have the properties that

$$\begin{cases} \boldsymbol{x}^* \in \arg \underset{\boldsymbol{x} \in X}{\text{minimum}}\ f(\boldsymbol{x}) + (\boldsymbol{\mu}^*)^{\mathrm{T}} \boldsymbol{g}(\boldsymbol{x}), \\ u_i^* g_i(\boldsymbol{x}^*) = 0, \quad i = 1, \ldots, m, \\ \boldsymbol{g}(\boldsymbol{x}^*) \leq \mathbf{0}^m, \end{cases}$$

then $\boldsymbol{x}^*$ is an optimal solution to (P). (In addition, $\boldsymbol{\mu}^*$ is an optimal solution to the Lagrangian dual problem.)

13.1 The Theorem

Let $\bar{\boldsymbol{\mu}} \geq \mathbf{0}^m$. Use the result above to prove that if $\bar{\boldsymbol{x}} \in X$ solves the Lagrangian relaxation to

$$\underset{\boldsymbol{x} \in X}{\text{minimize}}\ f(\boldsymbol{x}) + \bar{\boldsymbol{\mu}}^{\mathrm{T}} \boldsymbol{g}(\boldsymbol{x}),$$

then $\bar{\boldsymbol{x}}$ is an optimal solution to the problem to

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}), \\ \text{subject to} & g_i(\boldsymbol{x}) \leq \bar{y}_i, \quad i = 1, \ldots, m, \\ & \boldsymbol{x} \in X, \end{array}$$

where the original constraint right-hand sides have been perturbed into

$$\bar{y}_i \begin{cases} = g_i(\bar{\boldsymbol{x}}) \text{ for } \bar{\mu}_i > 0 \\ \geq g_i(\bar{\boldsymbol{x}}) \text{ for } \bar{\mu}_i = 0 \end{cases} \quad i = 1, \ldots, m.$$

(This result is known as *Everett's Theorem*.)

13.2 Numerical example

Consider the integer program to

$$\begin{array}{rrcrcrcrcrcrcl} \text{maximize } z = & 11x_1 & + & 6x_2 & + & 5x_3 & + & 15x_4 & + & 5x_5 & + & 4x_6 & & \\ \text{subject to} & 5x_1 & + & 3x_2 & + & 4x_3 & + & 6x_4 & + & 3x_5 & + & 4x_6 & \leq & b \\ & x_1 & + & x_2 & + & x_3 & & & & & & & = & 1 \\ & & & & & & & x_4 & + & x_5 & + & x_6 & = & 1 \\ & x_1 & , & x_2 & , & x_3 & , & x_4 & , & x_5 & , & x_6 & = & 0/1, \end{array}$$

where b is a parameter that can take on non-negative values. Lagrangian relax the knapsack constraint with the multiplier $\mu \geq 0$, and utilize Everett's Theorem to find an optimal solution for as many values of b as is possible. Can you, in this particular way, find the optimal solution for $b = 10$? (*Hint:* The Lagrangian dual objective function has break points at $\mu = 5/2$ and $\mu = 10/3$.)

13.3 Application

Consider the above integer program with $b = 8$. Suppose that the knapsack constraint is a *soft* resource constraint, that is, it does not need to be fulfilled exactly but may be violated somewhat. (This type of constraint is often encountered when there are uncertainties in the data, and when it is therefore not meaningful to require exact compliance with the constraint. It is also natural to model resource constraints in this way when it is possible to interpret a violation as a requirement for additional resource; the Lagrangian term then has the interpretation of an additional cost for this additional resource.) Suppose that we here allow for the constraint to be violated by a few units only. Use the above result to find a reasonable solution to this problem. (Many real-world optimization problems include constraints that in nature are soft; such constraints often models goals or resource limits. The opposite to such constraints are, of course, *hard* constraints. Such constraints must be fulfilled exactly, and often describe constitutional or logical conditions.)

[*Note:* According to Everett's Theorem an optimal solution to a Lagrangian relaxation of some problem is also a globally optimal solution to a version of the said problem in which the right-hand sides of the constraints have been perturbed. By varying the values of the multipliers used in the Lagrangian relaxation one can (normally) in addition obtain optimal solutions to primal problem with a large variety of perturbations of the right-hand sides. However, it is normally *not* possible to obtain *every* possible perturbation, and especially there normally do not exist multipliers that result in a zero perturbation (that is, $\bar{\boldsymbol{y}} = \mathbf{0}^m$), which would have meant that we have solved the original problem. (An exception is the case of convex problems with a strictly convex objective function: for such problems we always obtain the optimal solution to the original

problem when we use optimal multipliers in the Lagrangian subproblem.) The perturbations do, however, in some sense become smaller as the multiplier values gets closer to the optimal ones.

In summary, then, by solving Lagrangian subproblems (which normally is much less demanding than to solve the original problem) we find the exact optimal solution to *some* perturbed primal problems, but usually *not* that of the original problem.

Chapter 14

Surrogate relaxation

Consider a problem of the form

$$\begin{aligned} z^* = \ & \text{minimum} \ f(\boldsymbol{x}), \\ & \text{subject to} \ g_i(\boldsymbol{x}) \leq 0, \quad i = 1, \dots, m, \qquad (P) \\ & \boldsymbol{x} \in X, \end{aligned}$$

where $X \subseteq \mathbb{R}^n$ is closed and bounded, and the functions $f, g_i : \mathbb{R}^n \to \mathbb{R}$ are continuous. The problem is assumed to have an optimal solution, $\boldsymbol{x}^*$. Introduce parameters $\mu_i \geq 0$, $i = 1, \dots, m$, and define

$$\begin{aligned} s(\boldsymbol{\mu}) = \ & \text{minimum} \ f(\boldsymbol{x}), \\ & \text{subject to} \ \sum_{i=1}^{m} \mu_i g_i(\boldsymbol{x}) \leq 0 \qquad (S) \\ & \boldsymbol{x} \in X. \end{aligned}$$

This problem has only *one* explicit constraint.

14.1 Weak duality

Show that $\boldsymbol{x}^*$ is a feasible solution to the problem (S) and that therefore $s(\boldsymbol{\mu}) \leq z^*$ always holds, that is, (S) is a *relaxation* of the original problem. Motivate why $\text{maximum}_{\boldsymbol{\mu} \geq \boldsymbol{0}^m} \ s(\boldsymbol{\mu}) \leq z^*$ holds. Explain the potential use of this result!

14.2 Numerical example

Consider the linear 0/1 problem to find

$$\begin{array}{lllllllllll} z^* = & \text{maximum} \ z = & 5x_1 & + & 8x_2 & + & 7x_3 & + & 9x_4 & & & \\ & \text{subject to} & 3x_1 & + & 2x_2 & + & 3x_3 & + & 3x_4 & \leq & 6 & (1) \\ & & 2x_1 & + & 3x_2 & + & 3x_3 & + & 4x_4 & \leq & 5 & (2) \\ & & 2x_1 & + & x_2 & + & 2x_3 & + & x_4 & = & 3 & \\ & & x_1 & , & x_2 & , & x_3 & , & x_4 & \in & \{0,1\}. & \end{array}$$

Surrogate relax the constraints (1) and (2) with multipliers $\mu_1, \mu_2 \geq 0$ and state the problem (S). Let $\bar{\boldsymbol{\mu}} = (1,2)^{\mathrm{T}}$. Evaluate $s(\bar{\boldsymbol{\mu}})$.

Consider again the original problem, and Lagrangian relax the constraints (1) and (2) with multipliers $\mu_1, \mu_2 \geq 0$. Evaluate the Lagrangian dual function value at $\boldsymbol{\mu} = \bar{\boldsymbol{\mu}}$. Compare the two results.

14.3 Comparison with Lagrangian duality

Let $\boldsymbol{\mu} \geq \boldsymbol{0}^m$ and

$$q(\boldsymbol{\mu}) := \underset{\boldsymbol{x} \in X}{\text{minimum}} \ f(\boldsymbol{x}) + \sum_{i=1}^{m} \mu_i g_i(\boldsymbol{x}).$$

Show that $q(\boldsymbol{\mu}) \leq s(\boldsymbol{\mu})$, and that

$$\underset{\boldsymbol{\mu} \geq \boldsymbol{0}^m}{\text{maximum}} \ q(\boldsymbol{\mu}) \leq \underset{\boldsymbol{\mu} \geq \boldsymbol{0}^m}{\text{maximum}} \ s(\boldsymbol{\mu}) \leq z^*.$$

Part II

Column and constraint generation methods

Chapter 15

A cutting plane method for linear programs

Consider the standard LP problem to find

$$\begin{aligned} v_{LP} = \text{minimum} \quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}, \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}, \\ & \boldsymbol{D}\boldsymbol{x} \leq \boldsymbol{d}, \\ & \boldsymbol{x} \in \mathbb{R}^n_+. \end{aligned}$$

We suppose for now that X is bounded.

Let $P_X := \{\boldsymbol{x}^1, \boldsymbol{x}^2, \dots, \boldsymbol{x}^K\}$ be the set of extreme points in the polyhedron $X := \{\, \boldsymbol{x} \in \mathbb{R}^n_+ \mid \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\,\}$.

15.1 Reformulations

Its Lagrangian dual with respect to Lagrangian relaxing the constraints $\boldsymbol{D}\boldsymbol{x} \leq \boldsymbol{d}$ is to find

$$\begin{aligned} v_{LP} = v_L := \text{maximum} \quad & q(\boldsymbol{\mu}), \\ \text{subject to} \quad & \boldsymbol{\mu} \geq \boldsymbol{0}, \end{aligned}$$

where

$$\begin{aligned} q(\boldsymbol{\mu}) &:= \underset{\boldsymbol{x} \in X}{\text{minimum}} \left\{ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{x} - \boldsymbol{d}) \right\} \\ &= \underset{i \in P_X}{\text{minimum}} \left\{ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^i + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{x}^i - \boldsymbol{d}) \right\}. \end{aligned}$$

Show that this can equivalently be written as

$$q(\boldsymbol{\mu}) \leq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^i + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{x}^i - \boldsymbol{d}), \qquad i \in P_X, \quad \boldsymbol{\mu} \geq \boldsymbol{0}$$

that is,

$$\begin{aligned} v_L := \text{maximum} \quad & z, \\ \text{subject to} \quad & z \leq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^i + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{x}^i - \boldsymbol{d}), \qquad i \in P_X, \\ & \boldsymbol{\mu} \geq \boldsymbol{0}. \end{aligned}$$

15.2 A restriction; optimality

Suppose only a subset of P_X is known, and consider the following restriction of the Lagrangian dual problem:

$$z^{k+1} := \max \; z, \tag{15.1a}$$
$$\text{s.t.} \; z \leq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^i + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{x}^i - \boldsymbol{d}), \quad i = 1, \dots, k, \tag{15.1b}$$
$$\boldsymbol{\mu} \geq \boldsymbol{0}. \tag{15.1c}$$

How do we determine if we have found the optimal solution to the Lagrangian dual problem? Derive the optimality test! Suppose that we found the optimal solution. Explain how we obtain the primal optimal solution.

15.3 Progress

Suppose that we have not yet found the optimal solution to the Lagrangian dual problem. Explain how progress can be made based on the new information that was obtained from the optimality test.

Chapter 16

Dantzig–Wolfe decomposition

The cutting plane developed above has strong duality relationships with the classic Dantzig–Wolfe decomposition method for linear programming. The purpose of this exercise is to determine exactly which ones.

We rewrite the problem (15.1) as follows:

$$\begin{aligned} \underset{(z,\boldsymbol{\mu})}{\text{maximize}} \quad & z, \\ \text{subject to} \quad & z - \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{x}^i - \boldsymbol{d}) \leq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^i, \quad i = 1, \dots, k, \\ & \boldsymbol{\mu} \geq \boldsymbol{0}. \end{aligned}$$

With LP dual variables $\lambda_i \geq 0$ for the linear constraints, we obtain the LP dual to find

$$\begin{aligned} v^{k+1} = \text{minimum} \quad & \sum_{i=1}^{k} (\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^i)\lambda_i, \\ \text{subject to} \quad & \sum_{i=1}^{k} \lambda_i = 1, \\ & -\sum_{i=1}^{k} (\boldsymbol{D}\boldsymbol{x}^i - \boldsymbol{d})\lambda_i \geq \boldsymbol{0}, \\ & \lambda_i \geq 0, \qquad i = 1, \dots, k, \end{aligned}$$

that is,

$$\begin{aligned} v^{k+1} = \text{minimum} \quad & \boldsymbol{c}^{\mathrm{T}} \left(\sum_{i=1}^{k} \lambda_i \boldsymbol{x}^i \right), \\ \text{subject to} \quad & \sum_{i=1}^{k} \lambda_i = 1, \\ & \lambda_i \geq 0, \qquad i = 1, \dots, k, \\ & \boldsymbol{D} \left(\sum_{i=1}^{k} \lambda_i \boldsymbol{x}^i \right) \leq \boldsymbol{d}. \end{aligned} \tag{16.1}$$

16.1 Interpretations

Interpret the above problem in terms of the original, primal problem.

16.2 Duality relationships

The problem (16.1) is known as the *restricted master problem* (RMP) in the Dantzig–Wolfe algorithm. Explain the progress of this algorithm. How do we determine whether we have found an optimal solution? How do we obtain a primal optimal solution? Explain the exact duality relationships between the sub- and master problems in the cutting plane and Dantzig–Wolfe algorithms.

Chapter 17

Benders decomposition

Consider the following problem to

$$\begin{aligned} \text{minimize } & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + f(\boldsymbol{y}), \\ \text{subject to } & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{F}(\boldsymbol{y}) \geq \boldsymbol{b}, \\ & \boldsymbol{x} \geq \boldsymbol{0}^n, \quad \boldsymbol{y} \in S. \end{aligned}$$

We consider the variables $\boldsymbol{y}$ to be *difficult* in the sense that the problem in $\boldsymbol{x}$ is linear. The idea behind Benders decomposition is to utilize this property, through restricting the problem by fixing the vector $\boldsymbol{y}$ to a feasible value.

17.1 Applications

Provide an application of the above problem formulation.

17.2 Derivations, subproblem

We introduce the set

$$R := \{\, \boldsymbol{y} \in S \mid \exists \boldsymbol{x} \geq \boldsymbol{0}^n \text{ with } \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} - \boldsymbol{F}(\boldsymbol{y}) \,\}$$

What is the role of this set?

Further, we introduce the dual set

$$D = \{\, \boldsymbol{u} \in \mathbb{R}^m_+ \mid \boldsymbol{A}^{\mathrm{T}}\boldsymbol{u} \leq \boldsymbol{c} \,\}$$

together with the following entities:

$$\boldsymbol{u}_i^{\mathrm{p}}, \quad i = 1, \dots, n_p$$

and

$$\boldsymbol{u}_i^{\mathrm{r}}, \quad i = 1, \dots, n_r$$

denotes the extreme points and extreme rays of D, respectively.

Derive the Benders subproblem by using LP duality!

17.3 Derivations, master problem

Derive the restricted master problem!

17.4 Algorithm description

Describe the workings of the complete algorithm. In particular:

- How do we obtain a feasible solution to the problem?
- When do we know that we have found an optimal solution?
- Which are the requirements on the original problem data that will ensure convergence?
- In the case where the original problem is linear, what are the relationships between Benders decomposition and the cutting plane and Dantzig–Wolfe algorithms? What are your conclusions regarding the applicability of these three methods?
- Explain how these methods are best implemented in practice for a linear program.