TMA521/MMA510
Optimization, project course
Lecture 7
Cutting plane methods, column generation, and the Dantzig–Wolfe algorithm

Ann-Brith Strömberg

15 September 2010

A standard LP problem and its Lagrangian dual

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

- We suppose for now that the polyhedron $X := \{\, \mathbf{x} \in \mathbb{R}^n_+ \mid \mathbf{A}\mathbf{x} \leq \mathbf{b} \,\}$ is bounded
- Let $P_X := \{\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^K\}$ be the set of extreme points in X

The Lagrangian dual

- The Lagrangian dual of the LP with respect to relaxing the constraints $\mathbf{D}\mathbf{x} \leq \mathbf{d}$ is

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

- The Lagrangian dual function:

$$
q(\boldsymbol{\mu}) := \min_{\mathbf{x} \in X} \left\{ \mathbf{c}^{\mathrm{T}} \mathbf{x} + \boldsymbol{\mu}^{\mathrm{T}} (\mathbf{D}\mathbf{x} - \mathbf{d}) \right\} = \min_{i \in P_X} \left\{ \mathbf{c}^{\mathrm{T}} \mathbf{x}^i + \boldsymbol{\mu}^{\mathrm{T}} (\mathbf{D}\mathbf{x}^i - \mathbf{d}) \right\}
$$

- Solution set: $X(\boldsymbol{\mu}) := \arg\min_{\mathbf{x} \in X} \left\{ \mathbf{c}^{\mathrm{T}} \mathbf{x} + \boldsymbol{\mu}^{\mathrm{T}} (\mathbf{D}\mathbf{x} - \mathbf{d}) \right\}$
- Equivalent statement:

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

An equivalent formulation

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

- If, at an optimal dual solution $\boldsymbol{\mu}^*$, the solution set $X(\boldsymbol{\mu}^*)$ is a singleton, i.e., $X(\boldsymbol{\mu}^*) = \{\mathbf{x}^*\}$, then $\mathbf{x}^*$ is optimal (and unique) — thanks to strong duality
- This typically does not happen, unless an optimal solution $\mathbf{x}^*$ happens to be an extreme point of X
- But $\mathbf{x}^*$ can always be expressed as a convex combination of extreme points of X

A cutting plane method for the Lagrangian dual problem

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

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

- Let $(\boldsymbol{\mu}^k, z^k)$ be the solution to (1)
- It holds that $z^k \geq v_L$ for $k = 1, \ldots, K$
- How do we determine whether an optimal solution is found?
- And what IS the optimal solution when we find it?
- If $z^k \leq \mathbf{c}^{\mathrm{T}}\mathbf{x}^i + (\boldsymbol{\mu}^k)^{\mathrm{T}}(\mathbf{D}\mathbf{x}^i - \mathbf{d})$ holds for all $i \in P_X$, then $\boldsymbol{\mu}^k$ is optimal in the dual! Why?

Check optimality—generate new inequality

- How check optimality? Find the most violated constraint
- Solve the subproblem

$$q(\boldsymbol{\mu}^k) := \min_{\mathbf{x} \in X} \left\{ \mathbf{c}^{\mathrm{T}}\mathbf{x} + (\boldsymbol{\mu}^k)^{\mathrm{T}}(\mathbf{D}\mathbf{x} - \mathbf{d}) \right\} \tag{2}$$
$$= \min_{i \in P_X} \left\{ \mathbf{c}^{\mathrm{T}}\mathbf{x}^i + (\boldsymbol{\mu}^k)^{\mathrm{T}}(\mathbf{D}\mathbf{x}^i - \mathbf{d}) \right\}$$

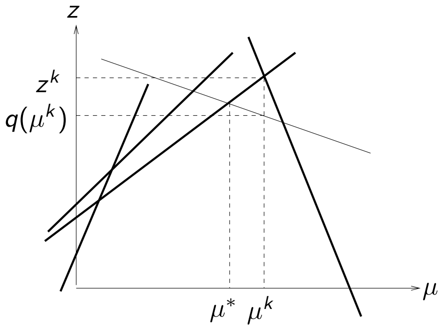
- If $z^k \leq q(\boldsymbol{\mu}^k)$ then $\boldsymbol{\mu}^k$ is optimal in the dual
- Otherwise, we have identified a constraint of the form

$$z \leq \mathbf{c}^{\mathrm{T}}\mathbf{x}^{k+1} + \boldsymbol{\mu}^{\mathrm{T}}(\mathbf{D}\mathbf{x}^{k+1} - \mathbf{d}),$$

  which is violated at $(\boldsymbol{\mu}^k, z^k)$ (i.e., it holds that $z^k > \mathbf{c}^{\mathrm{T}}\mathbf{x}^{k+1} + (\boldsymbol{\mu}^k)^{\mathrm{T}}(\mathbf{D}\mathbf{x}^{k+1} - \mathbf{d})$) f
- Add this inequality and re-solve the LP problem!

Cutting plane algorithm

- We call this a *cutting plane* algorithm
- It is based on adding constraints to the dual problem in order to improve the solution, in the process of cutting off the previous point
- Consider the below picture. The thick lines correspond to the subset of k inequalities known at iteration k



Cutting plane algorithm

- Obviously, $z^k \geq q(\mu^k)$ must hold, because of the possible lack of constraints.
- In this case, $z^k > q(\mu^k)$ holds, so in the next step when we evaluate $q(\mu^k)$ we can identify and add the last lacking inequality
- The resulting maximization will then yield the optimal solution μ^* shown in the picture
- How do we generate an optimal primal solution from this scheme? Let us look at the dual of the problem (1) in this cutting plane algorithm

Duality relations and the Dantzig–Wolfe algorithm

- We rewrite the problem (1)

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

- With LP dual variables $\lambda_i \geq 0$ we obtain the LP dual:

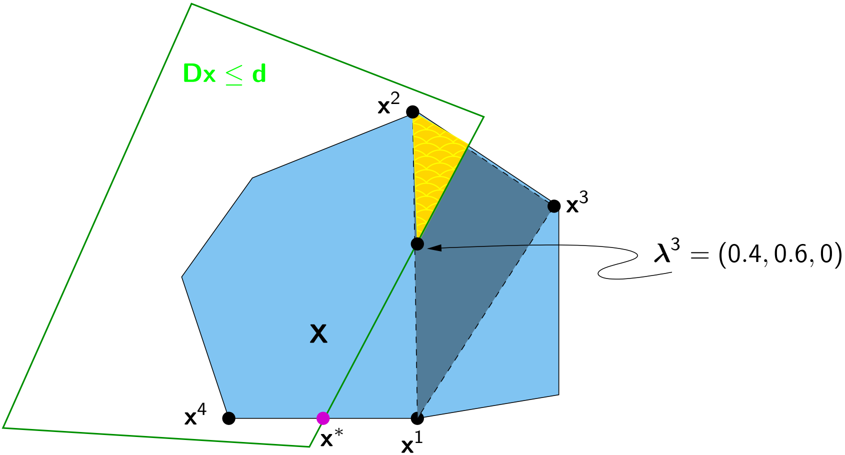
$$\begin{aligned} v^{k+1} = \text{minimum} \;\; & \sum_{i=1}^{k} (\mathbf{c}^{\mathrm{T}}\mathbf{x}^i)\lambda_i, \\ \text{subject to} \;\; & \sum_{i=1}^{k} \lambda_i = 1, \\ & -\sum_{i=1}^{k} (\mathbf{D}\mathbf{x}^i - \mathbf{d})\lambda_i \geq \mathbf{0}, \\ & \lambda_i \geq 0, \quad i = 1, \dots, k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

The linear programming dual rewritten

- $$v^{k+1} = \text{minimum } \mathbf{c}^{\mathrm{T}} \left(\sum_{i=1}^{k} \lambda_i \mathbf{x}^i \right), \tag{3}$$
$$\text{subject to} \qquad \sum_{i=1}^{k} \lambda_i = 1,$$
$$\lambda_i \geq 0, \qquad i = 1, \dots, k,$$
$$\mathbf{D} \left(\sum_{i=1}^{k} \lambda_i \mathbf{x}^i \right) \leq \mathbf{d}$$

- Maximize $\mathbf{c}^{\mathrm{T}}\mathbf{x}$ when $\mathbf{x}$ lies in the convex hull of the extreme points $\mathbf{x}^i$ found so far and fulfills the constraints that are Lagrangian relaxed

An illustration in x-space



The Dantzig-Wolfe algorithm

- The problem (3) is known as the restricted master problem (RMP) in the Dantzig–Wolfe algorithm
- In this algorithm, we have at hand a subset $\{1, \ldots, k\}$ of extreme points of X (and a dual vector $\boldsymbol{\mu}^{k-1}$)
- Find a feasible solution to the original LP problem by solving the restricted master problem (3)
- Then generate an optimal dual solution $\boldsymbol{\mu}^k$ to this restricted problem problem, corresponding to the constraints $\mathbf{Dx} \leq \mathbf{d}$
- If and only if the vector $\mathbf{x}^i$ generated in the next subproblem (2) was already included, we have found the optimal solution to the problem

Three algorithms which are "dual" to each other

- Cutting plane applied to the Lagrangian dual

$$\Longleftrightarrow$$

- Dantzig–Wolfe applied to the original LP

$$\Longleftrightarrow$$

- Benders decomposition applied to the dual LP.

Column generation

- Consider an LP with *very* many variables:
  $c_j, x_j \in \mathbb{R}$, $\mathbf{a}_j, \mathbf{b} \in \mathbb{R}^m$, $m \ll n$

$$
\begin{aligned}
\text{minimize} \quad & z = \sum_{j=1}^{n} c_j x_j \\
\text{subject to} \quad & \sum_{j=1}^{n} \mathbf{a}_j x_j = \mathbf{b} \\
& x_j \geq 0, \qquad j = 1, \ldots, n
\end{aligned}
$$

- The matrix $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ is too large to handle.
- Assume that m is relatively small $\Longrightarrow$ the basis matrix is not too large ($m \times m$)

Basic feasible solutions

- $B = \{m \text{ elements from the set } \{1, \dots, n\}\}$ is a basis if the corresponding matrix $\mathbf{B} = (\mathbf{a}_j)_{j \in B}$ has an inverse, $\mathbf{B}^{-1}$
- A basic solution is given by $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b}$ and $x_j = 0$, $j \notin B$. It is feasible if $\mathbf{x}_B \geq \mathbf{0}^m$
- A better basic feasible solution can be found by computing reduced costs: $\bar{c}_j = c_j - \mathbf{c}_B^{\mathrm{T}}\mathbf{B}^{-1}\mathbf{a}_j$ for $j \notin B$
- Let $\bar{c}_s = \underset{j \notin B}{\text{minimum}} \; \bar{c}_j$
- If $\bar{c}_s < 0 \Longrightarrow$ a better solution is received if x_s enters the basis
- If $\bar{c}_s \geq 0 \Longrightarrow \mathbf{x}_B$ is an optimal basic solution

Generating columns

- Suppose the columns $\mathbf{a}_j$ are defined by a set $S = \{\mathbf{a}_j \mid j = 1, \dots, n\}$ being, e.g., solutions to a system of equations (extreme points, integer points, ...)
- The incoming column is then chosen by solving a subproblem
$$\bar{c}(\mathbf{a}') = \operatorname*{minimum}_{\mathbf{a} \in S} \left\{ c - \mathbf{c}_B^{\mathrm{T}} \mathbf{B}^{-1} \mathbf{a} \right\}$$
- $\mathbf{a}'$ is a column having the least reduced cost w.r.t. the basis B
- If $\bar{c}(\mathbf{a}') < 0$ let the column $\begin{pmatrix} c(\mathbf{a}') \\ \mathbf{a}' \end{pmatrix}$ enter the problem

Example: The cutting stock problem

- **Supply:** rolls of e.g. paper of length L
- **Demand:** b_i roll pieces of length $\ell_i < L$, $i = 1, \ldots, m$
- **Objective:** minimize the number of rolls needed for producing the demanded pieces

First formulation

$$x_k = \begin{cases} 1 & \text{if roll } k \text{ is used} \\ 0 & \text{otherwise} \end{cases} \qquad y_{ik} = \begin{cases} 1 & \text{if piece } i \text{ is cut from roll } k \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} \text{minimize} \quad & \sum_{k=1}^{M} x_k \\ \text{subject to} \quad & \sum_{i=1}^{m} \ell_i y_{ik} \leq L x_k, \quad k = 1, \ldots, M \\ & \sum_{k=1}^{K} y_{ik} = b_i, \qquad i = 1, \ldots, m \\ & x_k, y_{ik} \text{ binary}, \quad i = 1, \ldots, m, k = 1, \ldots, M \end{aligned}$$

The value of the LP-relaxation is $\frac{\sum_{i=1}^{m} \ell_i b_i}{L}$ which can be very bad if $\ell_i = \lfloor L/2 + 1 \rfloor$ for large L
(large duality gap $\Rightarrow$ potentially bad performance of IP solvers)

Second formulation

- **Cut pattern:** number j contains a_{ij} pieces of length ℓ_i
- **Feasible** pattern if $\sum_{i=1}^{m} \ell_i a_{ij} \leq L$, where $a_{ij} \geq 0$, integer
- **Variables:** x_j = number of times pattern j is used

$$
\begin{aligned}
\text{minimize} \quad & \sum_{j=1}^{n} x_j \\
\text{subject to} \quad & \sum_{j=1}^{n} a_{ij} x_j = b_i, \qquad i = 1, \ldots, m \\
& x_j \geq 0, \ \text{integer}, \qquad j = 1, \ldots, n
\end{aligned}
$$

- **Bad news:** n = total number of feasible cut patterns—huge integer
- **Good news:** the value of the LP relaxation is often very close to that of the optimal solution.

$\Rightarrow$ Relax integrality constraints, solve an LP instead of an ILP

Starting solution

Trivial: m unit columns (gives lots of waste) $\Longrightarrow$

$$\begin{aligned} \text{minimize} \ & \sum_{j=1}^{m} x_j \\ \text{subject to} \ & x_j = b_j, \quad j = 1, \ldots, m \\ & x_j \geq 0, \quad j = 1, \ldots, m \end{aligned}$$

New columns

Generate better patterns using the dual variable values $\pi_i \implies$ new column

$$1 - \max_{a_{ik}} \sum_{i=1}^{m} \pi_i a_{ik} \qquad \left[\text{minimize} \left(c_k - \underbrace{\mathbf{c}_B^{\mathrm{T}} \mathbf{B}^{-1}}_{\pi} \mathbf{a}_k \right) \right]$$

$$\text{subject to} \sum_{i=1}^{m} \ell_i a_{ik} \leq L,$$

$$a_{ik} \geq 0, \text{ integer}, \qquad i = 1, \dots, m$$

Solution to this integer knapsack problem: new column $\mathbf{a}_k$