# Lecture 8: Cutting plane methods, column generation, and the Dantzig–Wolfe algorithm

Ann-Brith Strömberg

21 September 2009

# A standard LP problem and its Lagrangian dual

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

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

# The Lagrangian dual

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

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

where

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

- 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

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

- If, at an optimal dual solution $\boldsymbol{\mu}^*$, the solution set $X(\boldsymbol{\mu}^*)$ is a singleton, then—thanks to strong duality—this solution is optimal (and it is unique!).
- 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 written as a convex combination of such points.
- Let's see how it can be generated...

# A cutting plane method for the Lagrangian dual problem

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

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

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

# Check optimality—generate new inequality

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

$$\begin{aligned} q(\boldsymbol{\mu}^{k+1}) &:= \underset{\mathbf{x}\in X}{\text{minimum}} \left\{ \mathbf{c}^{\mathrm{T}}\mathbf{x} + (\boldsymbol{\mu}^{k+1})^{\mathrm{T}}(\mathbf{D}\mathbf{x} - \mathbf{d}) \right\} \qquad (2) \\ &= \underset{i\in P_X}{\text{minimum}} \left\{ \mathbf{c}^{\mathrm{T}}\mathbf{x}^i + (\boldsymbol{\mu}^{k+1})^{\mathrm{T}}(\mathbf{D}\mathbf{x}^i - \mathbf{d}) \right\}. \end{aligned}$$

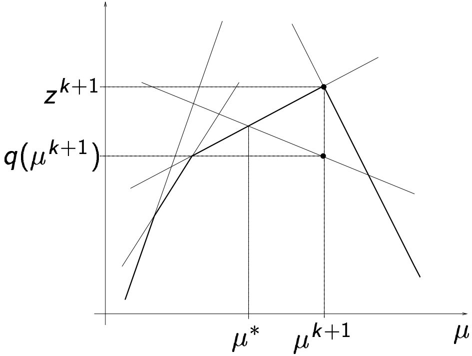
- If $z^{k+1} \leq q(\boldsymbol{\mu}^{k+1})$ then $\boldsymbol{\mu}^{k+1}$ is optimal in the dual; otherwise, we have identified a constraint of the form

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

which is violated at $(\boldsymbol{\mu}^{k+1}, z^{k+1})$.
Add this inequality and re-solve the LP problem!

# Cutting plane algorithm

- We call this a *cutting plane* algorithm, since it is based on adding constraints to the dual problem in order to improve the solution, in the process 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, $z^{k+1} \geq q(\mu^{k+1})$ must hold, because of the possible lack of constraints.
- In this case, $z^{k+1} > q(\mu^{k+1})$ holds, so in the next step when we evaluate $q(\mu^{k+1})$ we can identify and add the last lacking inequality
- The resulting maximization will then yield the optimal solution $\mu^*$ shown in the picture.
- What is the relationship to the standard simplex method?
- How do we generate a primal optimal solution from this scheme? Let us look at the dual of the problem (1) in this cutting plane algorithm.

- We rewrite the problem (1)

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

# The linear programming dual

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

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

# The linear programming dual rewritten

- Rewritten:

$$
\begin{aligned}
v^{k+1} = \text{minimum} \quad & \mathbf{c}^{\mathrm{T}} \left( \sum_{i=1}^{k} \lambda_i \mathbf{x}^i \right), && (3)\\
\text{subject to} \quad & \sum_{i=1}^{k} \lambda_i = 1, \\
& \lambda_i \geq 0, \qquad i = 1, \ldots, k, \\
& \mathbf{D} \left( \sum_{i=1}^{k} \lambda_i \mathbf{x}^i \right) \leq \mathbf{d}.
\end{aligned}
$$

- 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.

# 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 $\{1, \dots, k\}$ of extreme points of $X$ (and a dual vector $\boldsymbol{\mu}^k$).
- Find a feasible solution to the original LP problem by solving the restricted master problem (3).
- Then generate an optimal dual solution $\boldsymbol{\mu}^{k+1}$ 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:
  $c_j, x_j \in \mathbb{R}$, $\mathbf{a}_j, \mathbf{b} \in \mathbb{R}^m$, $m \ll n$

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

- The matrix $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ is too large to handle.
- Assume that $m$ is relatively small $\implies$ the basic 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 $\bar{c}_s < 0 \implies$ a better solution is received if $x_s$ enters the basis
- If $\bar{c}_s \geq 0 \implies \mathbf{x}_B$ is an optimal basic solution

# Generating columns

- Suppose the columns $\mathbf{a}_j$ are defined by a set $S = \{\mathbf{a}_j \mid j = 1, \ldots, n\}$ 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 $\ell_i$
- **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) $\Longrightarrow$

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

# New columns

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

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

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

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

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