Lecture 7: Distributed Control using Price Mechanisms

- **Motivation - why distributed optimization?**
- Duality in Linear Programming
- Finding optimum through distributed iterations

Distributed systems

Some systems are distributed by nature: e.g. Energy Production:

- Electric companies buy electricity from different producers, sell to consumers
- Produced electricity must match consumed electricity (which varies over day)
- Need reliable (but slow) source (e.g. nuclear power) and less reliable (but fast) source (e.g. wind power)
- How much reliable and how much fast electricity is needed?
- How much to pay for different electricity sources to get the desired amounts of electricity?

Distributed systems cont'd

Large production companies:

- Several sub-divisions, each producing several products
- Objective: Maximize company profit (not sub-division profit)
- Few common resource (e.g. packing) is shared

Can be optimized centrally by head-quarter. The resulting problem might be too large.

Distributed optimization:

- Each sub-division maximizes their profit (smaller problem)
- Head-quarter coordinates such that the common resource is fully used (if needed) and that the most profitable products are produced if common resource is limiting

Lecture 8: Distributed Optimization

- Motivation - why distributed optimization?
- **Duality in Linear Programming**
- Finding optimum through distributed iterations

Linear Programming Example

The following example is used throughout this lecture:

> A company consists of two sub-divisions. One sub-division manufactures garden furniture (by sawing and assembling), the other sub-division manufactures sleds (by sawing and assembling). Each division manufactures two different kinds of their respective products. Both sub-divisions send their products to a common painting station. The objective is to maximize company profit.

Linear Programming Example

Product	# of items	Profit / item
Garden Furniture 1	x_1	c_1
Garden Furniture 2	x_2	c_2
Sled 1	x_3	c_3
Sled 2	x_4	c_4

Constraints for sub-division 1:

$$7x_1 + 10x_2 \le 100 \qquad \text{(Sawing)}$$
$$16x_1 + 12x_2 \le 135 \qquad \text{(Assembling)}$$

Constraints for sub-division 2:

$$10x_3 + 9x_4 \le 70 \qquad \text{(Sawing)}$$
$$6x_3 + 9x_4 \le 60 \qquad \text{(Assembling)}$$

Painting Constraint:

$$5x_1 + 3x_2 + 3x_3 + 2x_4 \le 45$$

Linear Programming Example

Mathematical formulation:

$$\begin{array}{ll} \text{Maximize} & c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 \\ \text{subject to} & 7x_1 + 10x_2 \le 100 \\ & 16x_1 + 12x_2 \le 135 \\ & 10x_3 + 9x_4 \le 70 \\ & 6x_3 + 9x_4 \le 60 \\ & 5x_1 + 3x_2 + 3x_3 + 2x_4 \le 45 \\ & x \succeq 0 \end{array}$$

Linear Programming

The problem is on the following general form:

$$\begin{array}{ll} \text{Maximize} & c^T x \\ \text{subject to} & Ax \preceq b,\ x \succeq 0 \end{array}$$

which was studied in the MPC-lecture.

Over the next couple of slides we introduce the dual of this problem

Linear Programming Duality

Linear Program:

$$p^* = \begin{cases} \max\limits_x & c^T x \\ \text{subject to} & Ax \preceq b,\ x \succeq 0 \end{cases}$$

where $p^* = c^T x^*$ is the optimal value attained by x^*.

For the constraints $Ax \preceq b$, introduce dual variables $\lambda \succeq 0$ and construct the corresponding dual function $g(\lambda)$:

$$g(\lambda) = \max_{x \succeq 0} \left[c^T x + \lambda^T (b - Ax) \right]$$

The second term in the bracket is non-negative when $Ax \preceq b$. Hence $g(\lambda) \geq p^*$.

Linear Programming Duality cont'd

Tightest upper bound to p^* obtained by minimizing $g(\lambda)$:

$$d^* = \min_{\lambda \succeq 0} g(\lambda) = \min_{\lambda \succeq 0} \max_{x \succeq 0} \left[c^T x + \lambda^T (b - Ax) \right]$$

Optimal value d^* for this min-max problem is attained by $x = x^*$ and $\lambda = \lambda^*$.

Further we have that $p^* = c^T x^* = d^*$. This equality is referred to as *strong duality*

This min-max problem is used later to distribute the optimization

Dual optimal values and d^* can be obtained by solving

$$\begin{array}{ll} \min\limits_\lambda & b^T \lambda \\ \text{subject to} & A^T \lambda \succeq c, \lambda \succeq 0 \end{array}$$

Note symmetry to primal problem

Optimality Conditions

x^* is primal optimal if and only if there are dual variables λ^* such that

$$\begin{array}{rr} Ax^* \preceq b & A^T \lambda^* \succeq c \\ \lambda^* \succeq 0 & x^* \succeq 0 \\ (A_i x^* - b_i)\lambda_i^* = 0 & (A_j^T \lambda^* - c_j) x_j^* = 0 \end{array}$$

These conditions are called the KKT-conditions for this LP-problem

Interpretation of Dual Variables

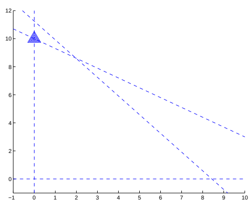
Dual variables can be interpreted as marginal price for resources:

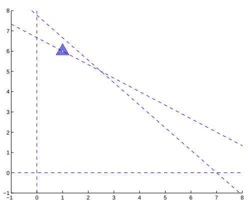
If the capacity for a resource is increased by 1, the total profit is increased by the corresponding dual variable.

This gives insight to which resource to increase to gain most

Numerical Results

Optimal solution for Division 1 (left) and Division 2 (right). Common constraint active (i.e. equality holds).





Numerical Results

Optimal dual variables and their respective constraints:

Constraint	Dual variable
$7x_1 + 10x_2 \leq 100$	1.04
$16x_1 + 12x_2 \leq 135$	0
$10x_3 + 9x_4 \leq 70$	0
$6x_3 + 9x_4 \leq 60$	0.4
$5x_1 + 3x_2 + 3x_3 + 2x_4 \leq 45$	3.2

Optimal value: $p^* = c^T x^* = 272$

If common (painting) constraint capacity increased to 46, optimal value becomes $272 + 3.2 = 275.2$

Company would gain most by increasing painting capacity

Lecture 8: Distributed Optimization

- Motivation - why distributed optimization?
- Duality in Linear Programming
- **Finding optimum through distributed iterations**

Distribution of LP Example

Solve the LP example

$$\begin{array}{ll} \text{Maximize} & c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 \\ \text{subject to} & 7x_1 + 10x_2 \leq 100 \\ & 16x_1 + 12x_2 \leq 135 \\ & 10x_3 + 9x_4 \leq 70 \\ & 6x_3 + 9x_4 \leq 60 \\ & 5x_1 + 3x_2 + 3x_3 + 2x_4 \leq 45 \\ & x \succeq 0 \end{array}$$

in a distributed fashion using the dual problem

Distribution of LP Example cont'd

Dual problem when constraint with all variables is "dualized":

$$\begin{array}{ll} \min\limits_{\lambda\geq 0}\max\limits_{x\geq 0} & c^T x + \lambda(45 - 5x_1 + 3x_2 + 3x_3 + 2x_4) \\ \text{subject to} & 7x_1 + 10x_2 \leq 100 \\ & 16x_1 + 12x_2 \leq 135 \\ & 10x_3 + 9x_4 \leq 70 \\ & 6x_3 + 9x_4 \leq 60 \end{array}$$

For fixed $\lambda = \bar\lambda$, the inner maximization can be decomposed to two sub-problems (one for each sub-division) P_1 and P_2:

$$P1: \left\{ \begin{array}{ll} \max\limits_{x_1\geq 0, x_2\geq 0} & c_1x_1 + c_2x_2 - \bar\lambda(5x_1 + 3x_2) \\ \text{s. t.} & 7x_1 + 10x_2 \leq 100 \\ & 16x_1 + 12x_2 \leq 135 \end{array}\right.$$

$$P2: \left\{ \begin{array}{ll} \max\limits_{x_3\geq 0, x_4\geq 0} & c_3x_3 + c_4x_4 - \bar\lambda(3x_3 + 2x_4) \\ \text{s. t.} & 10x_3 + 9x_4 \leq 70 \\ & 6x_3 + 9x_4 \leq 60 \end{array}\right.$$

Distribution Example cont'd

With fixed $x = \bar x$ head-quarters can update the dual variable λ to decrease the value of the outer minimization problem:

$$\bar\lambda^+ = \bar\lambda - \alpha(45 - 5\bar x_1 + 3\bar x_2 + 3\bar x_3 + 2\bar x_4)$$

where α is the step-size, which is chosen so that $\bar\lambda^+ \geq 0$ is maintained.

Motivation, the dual objective with $\bar\lambda$ is

$$g(\bar\lambda) = p^T\bar x + \bar\lambda(45 - 5\bar x_1 + 3\bar x_2 + 3\bar x_3 + 2\bar x_4)$$

and with $\bar\lambda^+$:

$$\begin{aligned} g(\bar\lambda^+) &= p^T\bar x + \bar\lambda^+(45 - 5\bar x_1 + 3\bar x_2 + 3\bar x_3 + 2\bar x_4) = \\ &= p^T\bar x + \bar\lambda(45 - 5\bar x_1 + 3\bar x_2 + 3\bar x_3 + 2\bar x_4) - \\ &\quad - \alpha(45 - 5\bar x_1 + 3\bar x_2 + 3\bar x_3 + 2\bar x_4)^2 \leq g(\bar\lambda) \end{aligned}$$

Distributed Optimization Algorithm

1. Initialize algorithm by $\lambda^{(0)} = 0$ and $x^{(0)} = 0$.
2. For fixed $\lambda = \lambda^{(k)}$ let the sub-divisions solve their respective optimization problems to find the state vector $x^{(k)}$.
3. Define $\lambda^{(k+1)} = \max(0, \lambda^{(k)} - \alpha^{(k)}(45 - 5x_1^{(k)} + 3x_2^{(k)} + 3x_3^{(k)} + 2x_4^{(k)}))$
4. Set $k \leftarrow k + 1$ and go to step 2.

Convergence to optimal value and convergence in dual variables guaranteed with this algorithm, if the step size λ^k is appropriately chosen

Convergence in primal variables guaranteed if objective strictly concave

A Convergence Theorem

Suppose $\|\lambda^{(1)} - \lambda^*\| \leq R$ and consider the iteration

$$\lambda^{(k+1)} = \lambda^{(k)} - \alpha_k g^{(k)}$$

where $g^{(k)}$ satisfies the "subgradient" inequality

$$f(\lambda^*) \geq f(\lambda^{(k)}) + (g^{(k)})^T(\lambda^* - \lambda^{(k)}) \qquad \text{for all } \lambda^{(k)}$$

and f satisfies the Lipschitz condition

$$|f(u) - f(v)| \leq G\|u - v\| \qquad \text{for all } u, v$$

Define $f_{\text{best}}^{(k)} = \min\{f(\lambda^{(1)}), \ldots, f(\lambda^{(k)})\}$. Then

$$f_{\text{best}}^{(k)} - f(\lambda^*) \leq \frac{R^2 + G^2\sum_{i=1}^k \alpha_k^2}{2\sum_{i=1}^k \alpha_k}$$

In particular $f_{\text{best}}^{(k)} \to f(\lambda^*)$ as $k \to \infty$ if $\alpha_k = \frac{1}{k}$.

Distributed Optimization Algorithm

Problem: Objective in LP not strictly concave (only concave)

This issue can be resolved in different ways:

- Add concave term to objective that do not alter optimal solution (might be difficult to keep distributed structure of the problem)
- A convex combination of all generated primal variables $\sum_{j=0}^{k} \mu_j^k x^j$ with certain requirements on μ_j^k and on the step-parameter α^k gives convergence in primal variables

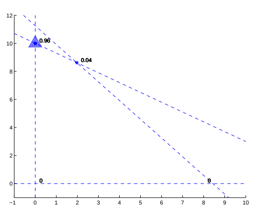
This is not pursued further here

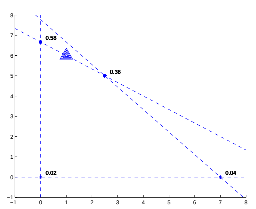
Comments on Distributed Optimization

- Decomposition scheme is called dual decomposition
- Dual decomposition most useful for large problems with
  - few constraints involving all variables
  - many local constraints
- Applicable to other types of optimization problems as well (such as quadratic problems)

Numerical Results

Primal variable iterates (x) for division 1 (left) and division 2 (right) with their respective local constraints. Triangles show optimal solution (which is not in a corner in division 2 due to the constraint with all variables). The numbers show the fraction of iterates in that corner.





Numerical Results

Same as previous slide where a certain convex combination of the solutions is plotted. These converge to the primal optimal solution. The numbers correspond to iterate number.