Distributed Optimization

Lecture 3: (Augmented) Lagrangian Relaxation

Jalal Kazempour

June 21, 2019
Learning objectives

After Lecture 3, you are expected to be able to:

- Explain the functioning of Lagrangian relaxation (LR), augmented Lagrangian relaxation (ALR), and alternating direction method of multipliers (ADMM)
- Implement them to an illustrative example
Decomposition techniques

Applicable to optimization problems with complicating constraints:

• Lagrangian relaxation (LR)
  
  In the literature, this technique has also been known as standard or conventional LR (or dual decomposition)!

• Augmented Lagrangian relaxation (ALR)
  
  - Auxiliary problem principle (APP)
  
  - Alternating direction method of multipliers (ADMM)

• Dantzig-Wolfe decomposition

• ...

In the literature, this technique has also been known as standard or conventional LR (or dual decomposition)!
Decomposition techniques

Applicable to optimization problems with complicating constraints:

- Lagrangian relaxation (LR)
  In the literature, this technique has also been known as *standard* or *conventional* LR (or *dual decomposition*)!

- Augmented Lagrangian relaxation (ALR)
  - Auxiliary problem principle (APP)
  - Alternating direction method of multipliers (ADMM)

- Dantzig-Wolfe decomposition

- ...

Will not be covered in this course!
Potential applications in power systems
Potential applications in power systems

• Market-clearing problem
  ✓ **Complicating constraints**: balance equalities and ramping limits of generators
  ✓ If relaxed, the original problem decomposes by agent (and by hour)
Potential applications in power systems

• Market-clearing problem

  ✓ **Complicating constraints**: balance equalities and ramping limits of generators
  ✓ If relaxed, the original problem decomposes by agent (and by hour)

• Multi-regional market-clearing (or unit commitment) problem, e.g., in case of pan-European electricity market

  ✓ **Complicating constraints**: tie-line constraints (power flow, and tie-line capacity)
  ✓ If relaxed, the original problem decomposes by region. This way, the operator of each region only solves its own market-clearing problem (the so-called distributed market-clearing problem)
Main references


Lagrangian relaxation (LR)

Background:

• The theory of LR (and also augmented LR) was firstly developed for problems with continuous variables, and functions (objective function and constraints) with first derivatives continuous.
Lagrangian relaxation (LR)

Background:

• The theory of LR (and also augmented LR) was firstly developed for problems with *continuous* variables, and functions (objective function and constraints) with *first* derivatives continuous.

• However, the theory has been used in problems with *binary* variables (like unit commitment problems) with success.
Lagrangian relaxation (LR)

*Background:*

- LR works efficiently if the number of complicating constraints is relatively low, and it is OK to have binary variables in the formulation.

- LR was extensively used in the 90’s to solve unit commitment problems (complicating constrains are just balance constraints and ramping constraints).
Lagrangian relaxation (LR)

Key point

In case of LR:

In addition to convexity, the objective function of the original (non-decomposed) problem needs to be smooth (continuous first derivatives), e.g., quadratic. If this objective function is linear, the LR procedure does not necessarily converge!
Lagrangian relaxation (LR)

Key point

In case of LR:

In addition to convexity, the objective function of the original (non-decomposed) problem needs to be smooth (continuous first derivatives), e.g., quadratic. If this objective function is linear, the LR procedure does not necessarily converge!

• Alternative solution technique for problems with linear objective function is augmented LR.
Lagrangian relaxation (LR)

For unit commitment (and also market clearing) problems:

- LR (for problems with quadratic objective function)
- ALR (for problems with either quadratic or linear objective function)

Both have been extensively and very successfully used in the literature, though unit commitment problem is non-convex (due to binary variables).
LR: Mathematical procedure

Consider the following optimization problem:

Minimize \( \sum_{i=1}^{l} f_i(x_i) \)

Subject to

\( g_i(x_i) = A_i \quad \forall i \)
\( h_i(x_i) \leq B_i \quad \forall i \)
\( \sum_{i=1}^{l} c_i(x_i) = M \quad (\lambda) \)
\( \sum_{i=1}^{l} d_i(x_i) \leq N \quad (\mu) \)
LR: Mathematical procedure

Consider the following optimization problem:

Minimize \( \sum_{i=1}^{l} f_i(x_i) \)

Subject to

\[ g_i(x_i) = A_i \quad \forall i \]
\[ h_i(x_i) \leq B_i \quad \forall i \]
\[ \sum_{i=1}^{l} c_i(x_i) = M \quad (\lambda) \]
\[ \sum_{i=1}^{l} d_i(x_i) \leq N \quad (\mu) \]

Is it decomposable?
LR: Mathematical procedure

Consider the following optimization problem:

Minimize \[ \sum_{i=1}^{l} f_i(x_i) \]

Subject to
\[ g_i(x_i) = A_i \quad \forall i \]
\[ h_i(x_i) \leq B_i \quad \forall i \]
\[ \sum_{i=1}^{l} c_i(x_i) = M \quad (\lambda) \]
\[ \sum_{i=1}^{l} d_i(x_i) \leq N \quad (\mu) \]

Complicating constraints
Consider the following optimization problem:

Minimize \[ \sum_{i=1}^{l} f_i(x_i) \]

Subject to

\[ g_i(x_i) = A_i \quad \forall i \]
\[ h_i(x_i) \leq B_i \quad \forall i \]
\[ \sum_{i=1}^{l} c_i(x_i) = M \quad (\lambda) \]
\[ \sum_{i=1}^{l} d_i(x_i) \leq N \quad (\mu) \]

Complicating constraints

Dual variables
(Lagrangian multipliers)
LR: Mathematical procedure

The original problem is equivalent to its Lagrangian dual problem (a max-min problem):

\[
\begin{align*}
\text{Maximize}_{\lambda, \mu} & \quad \sum_{i=1}^{l} f_i(x_i) + \lambda \left( M - \sum_{i=1}^{l} c_i(x_i) \right) + \mu \left( N - \sum_{i=1}^{l} d_i(x_i) \right) \\
\text{Subject to} & \quad g_i(x_i) = A_i \quad \forall i \\
& \quad h_i(x_i) \leq B_i \quad \forall i
\end{align*}
\]
LR: Mathematical procedure

The original problem is equivalent to its Lagrangian dual problem (a max-min problem):

\[
\begin{align*}
\text{Maximize} & \quad \lambda, \mu \\
\text{Minimize} & \quad x_i \sum_{i=1}^{I} f_i(x_i) + \lambda \left[ M - \sum_{i=1}^{I} c_i(x_i) \right] + \mu \left[ N - \sum_{i=1}^{I} d_i(x_i) \right] \\
\text{Subject to} & \quad g_i(x_i) = A_i \quad \forall i \\
& \quad h_i(x_i) \leq B_i \quad \forall i
\end{align*}
\]

Is this equivalent problem decomposed?
The original problem is equivalent to its *Lagrangian dual problem* (a max-min problem):

\[
\begin{align*}
\text{Maximize } & \quad \lambda, \mu \\
\text{Minimize } & \quad \sum_{i=1}^{l} f_i(x_i) + \lambda \left[ M - \sum_{i=1}^{l} c_i(x_i) \right] + \mu \left[ N - \sum_{i=1}^{l} d_i(x_i) \right] \\
\text{Subject to } & \quad g_i(x_i) = A_i \quad \forall i \\
& \quad h_i(x_i) \leq B_i \quad \forall i
\end{align*}
\]

Is this equivalent problem decomposed? **Not yet!**
LR: Mathematical procedure

The original problem is equivalent to its *Lagrangian dual problem* (a max-min problem):

\[
\begin{align*}
\text{Maximize} & \quad \lambda, \mu \\
\text{Minimize} & \quad \sum_{i=1}^{I} f_i(x_i) + \lambda \left[ M - \sum_{i=1}^{I} c_i(x_i) \right] + \mu \left[ N - \sum_{i=1}^{I} d_i(x_i) \right] \\
\text{Subject to} & \\
& g_i(x_i) = A_i \quad \forall i \\
& h_i(x_i) \leq B_i \quad \forall i
\end{align*}
\]

Is this equivalent problem decomposed? **Not yet!**

- Let's *relax* the equivalent problem above by *fixing dual variables* ($\lambda$ and $\mu$) to given values, i.e., $\bar{\lambda}$ and $\bar{\mu}$.
LR: Mathematical procedure

The relaxed problem:

Minimize \( \sum_{i=1}^{l} f_i(x_i) + \bar{\lambda} \left[ M - \sum_{i=1}^{l} c_i(x_i) \right] + \bar{\mu} \left[ N - \sum_{i=1}^{l} d_i(x_i) \right] \)

Subject to

\( g_i(x_i) = A_i \quad \forall i \)
\( h_i(x_i) \leq B_i \quad \forall i \)
LR: Mathematical procedure

The relaxed problem:

Minimize \( \sum_{i=1}^{l} f_i(x_i) + \bar{\lambda} \left[ M - \sum_{i=1}^{l} c_i(x_i) \right] + \bar{\mu} \left[ N - \sum_{i=1}^{l} d_i(x_i) \right] \)

Subject to

\( g_i(x_i) = A_i \quad \forall i \)
\( h_i(x_i) \leq B_i \quad \forall i \)

Parameters (fixed values)
LR: Mathematical procedure

The relaxed problem:

Minimize \[ \sum_{i=1}^{l} f_i(x_i) + \bar{\lambda} \left[ M - \sum_{i=1}^{l} c_i(x_i) \right] + \bar{\mu} \left[ N - \sum_{i=1}^{l} d_i(x_i) \right] \]

Subject to
\[ g_i(x_i) = A_i \quad \forall i \]
\[ h_i(x_i) \leq B_i \quad \forall i \]

Is this relaxed problem decomposable?
The relaxed problem:

Minimize \[ \sum_{i=1}^{I} f_i(x_i) + \lambda \left[ M - \sum_{i=1}^{I} c_i(x_i) \right] + \mu \left[ N - \sum_{i=1}^{I} d_i(x_i) \right] \]

Subject to
g_i(x_i) = A_i \quad \forall i
h_i(x_i) \leq B_i \quad \forall i

Is this relaxed problem decomposable? Yes, one per \( i \):

\[
\begin{aligned}
\text{Minimize} & \quad f_i(x_i) + \lambda c_i(x_i) + \mu d_i(x_i) \\
\text{Subject to} & \quad g_i(x_i) = A_i \\
& \quad h_i(x_i) \leq B_i 
\end{aligned}
\]
LR: Mathematical procedure

LR is an iterative approach with a systematic way to update the values of fixed dual variables ($\tilde{\lambda}$ and $\tilde{\mu}$) in each iteration.
**LR: Mathematical procedure**

LR is an iterative approach with a systematic way to update the values of fixed dual variables ($\bar{\lambda}$ and $\bar{\mu}$) in each iteration.

Available techniques in the literature to update $\bar{\lambda}$ and $\bar{\mu}$ [1]:

1. Subgradient method
2. Cutting plane method
3. Bundle method
4. Trust region method
5. ...

LR: Mathematical procedure

LR is an iterative approach with a systematic way to update the values of fixed dual variables ($\tilde{\lambda}$ and $\tilde{\mu}$) in each iteration.

Available techniques in the literature to update $\tilde{\lambda}$ and $\tilde{\mu}$ [1]:

1. Subgradient method
2. Cutting plane method
3. Bundle method
4. Trust region method
5. ...

Will not be covered in this course!

LR: illustrative example

Minimize $x^2 + y^2$

$x \geq 0, y \geq 0$

Subject to $-x - y = -4 \quad (\mu)$

Note: Objective function includes quadratic terms, so LR works!
LR: illustrative example

Minimize $x^2 + y^2$
\[x \geq 0, y \geq 0\]
Subject to $-x - y = -4$ \ ($\mu$)

Note: Objective function includes quadratic terms, so LR works!

Subproblem 1:

Minimize $x^2 - \bar{\mu}x$
\[x \geq 0\]

Subproblem 2:

Minimize $y^2 - \bar{\mu}y$
\[y \geq 0\]
LR: illustrative example

Minimize \( x^2 + y^2 \)
\[ x \geq 0, y \geq 0 \]

Subject to \(-x - y = -4 \) \((\mu)\)

Note: Objective function includes quadratic terms, so LR works!

Subproblem 1:
Minimize \( x^2 - \bar{\mu}x \)
\[ x \geq 0 \]

Subproblem 2:
Minimize \( y^2 - \bar{\mu}y \)
\[ y \geq 0 \]

Updating fixed dual variable \((\bar{\mu})\) using subgradient method:

- Solve subproblems 1 and 2 in iteration \(v\), and obtain the values \(x^{(v)}\) and \(y^{(v)}\)
- \(\bar{\mu}^{(v+1)} \leftarrow \bar{\mu}^{(v)} + \frac{1}{a + bv} \frac{-x^{(v)} - y^{(v)} + 4}{[-x^{(v)} - y^{(v)} + 4]}\)
- \(a\) and \(b\) are positive constants, e.g., \(a = 1\) and \(b = 0.1\).
Algorithm:

- **Step 0: Initialization**
  
  \[ \mu = 1 \text{ and } \mu^{(1)} = \mu^{\text{initial}} \]

- **Step 1: Solve subproblems 1 and 2, and obtain \( x^{(v)} \) and \( y^{(v)} \)

- **Step 2: Update fixed dual variable, i.e., \( \mu^{(v+1)} \)

- **Step 3: Convergence check**
  
  If \[ \frac{\|\mu^{(v+1)} - \mu^{(v)}\|}{\|\mu^{(v)}\|} \leq \epsilon \], then the optimal solution with a level of accuracy \( \epsilon \) is obtained, otherwise \( v \leftarrow v + 1 \) and go Step 1
Optional Assignment 1

Provide a GAMS (or Python or Julia) code for the previous illustrative example solved by LR algorithm!
Augmented Lagrangian relaxation (ALR)

Recall:

ALR works for problems with either quadratic objective function (like LR) or linear one (unlike LR)

Main difference of ALR with respect to LR:

An additional penalty term within the subproblems
Augmented Lagrangian relaxation (ALR)

Recall the previous illustrative example:

\[
\begin{align*}
\text{Minimize } & \quad x^2 + y^2 \\
\text{subject to } & \quad -x - y = -4 \quad (\lambda)
\end{align*}
\]
Augmented Lagrangian relaxation (ALR)

Recall the previous illustrative example:

\[
\begin{align*}
\text{Minimize } & \quad x^2 + y^2 \\
\text{Subject to } & \quad -x - y = -4 \quad (\lambda)
\end{align*}
\]

Equivalent to:

\[
\begin{align*}
\text{Maximize } & \quad \text{Minimize } x^2 + y^2 + \lambda(-x - y + 4) \\
\text{Minimize } & \quad x^2 + y^2 \quad \lambda(x,y)
\end{align*}
\]
Augmented Lagrangian relaxation (ALR)

Recall the previous illustrative example:

\[
\begin{align*}
\text{Minimize } & \quad x^2 + y^2 \\
\text{subject to } & \quad -x - y = -4 \quad (\lambda)
\end{align*}
\]

Equivalent to:

\[
\begin{align*}
\text{Maximize Minimize } & \quad x^2 + y^2 + \lambda(-x - y + 4) + \frac{\gamma}{2} \| -x - y + 4 \|^2
\end{align*}
\]

Additional penalty term with respect to LR, whose value is zero in the optimal point. \(\gamma\) is a positive constant.
Augmented Lagrangian relaxation (ALR)

Recall the previous illustrative example:

\[
\text{Minimize } x^2 + y^2 \\
\text{subject to } -x - y = -4 (\lambda)
\]

Equivalent to:

\[
\max_{\lambda} \min_{x \geq 0, y \geq 0} x^2 + y^2 + \lambda(-x - y + 4) + \frac{\gamma}{2} \| -x - y + 4 \|^2
\]

Additional penalty term with respect to LR, whose value is zero in the optimal point. \( \gamma \) is a positive constant.

Since it is a quadratic term, the first derivatives of the objective function with respect to variables are now continuous (not fixed). This is necessary for convergence.
Augmented Lagrangian relaxation (ALR)

Recall the previous illustrative example:

\[
\text{Minimize } x^2 + y^2 \\
\text{Subject to } -x - y = -4 \quad (\lambda)
\]

Equivalent to:

\[
\text{Maximize Minimize } x^2 + y^2 + \lambda(-x - y + 4) + \frac{\gamma}{2} \| -x - y + 4 \|^2
\]

Additional penalty term with respect to LR, whose value is zero in the optimal point.
\(\gamma\) is a positive constant

Question:

• Similar to LR, assume dual variable \(\lambda\) is fixed to a given value \(\bar{\lambda}\). Is the problem above decomposed for given \(\bar{\lambda}\)?
Augmented Lagrangian relaxation (ALR)

Recall the previous illustrative example:

\[
\begin{align*}
\text{Minimize} & \quad x^2 + y^2 \\
\text{Subject to} & \quad -x - y = -4 \quad (\lambda)
\end{align*}
\]

Equivalent to:

\[
\begin{align*}
\text{Maximize Minimize} & \quad x^2 + y^2 + \lambda(-x - y + 4) + \frac{\gamma}{2} \| -x - y + 4 \| ^2
\end{align*}
\]

Question:

- Similar to LR, assume dual variable $\lambda$ is fixed to a given value $\bar{\lambda}$. Is the problem above decomposed for given $\bar{\lambda}$? No, due to product of $x$ and $y$ in the penalty term!

Additional penalty term with respect to LR, whose value is zero in the optimal point. $\gamma$ is a positive constant.
Augmented Lagrangian relaxation (ALR)

Available alternatives to solve ALR:

• Auxiliary problem principle (APP)

• Alternating direction method of multipliers (ADMM)
Augmented Lagrangian relaxation (ALR)

Available alternatives to solve ALR:

• Auxiliary problem principle (APP): will not be covered in this course

• Alternating direction method of multipliers (ADMM)
Augmented Lagrangian relaxation (ALR)

Available alternatives to solve ALR:

• Auxiliary problem principle (APP): *will not be covered in this course*

• Alternating direction method of multipliers (ADMM)

  - ADMM directly **fixes** each variable to its value obtained in the previous iteration, and decomposes the ALR to subproblems.
Augmented Lagrangian relaxation (ALR)

Available alternatives to solve ALR:

• Auxiliary problem principle (APP): *will not be covered in this course*

• Alternating direction method of multipliers (ADMM)

  ❑ ADMM directly fixes each variable to its value obtained in the previous iteration, and decomposes the ALR to subproblems.

  ❑ Proof for convergence of ADMM to the optimal solution (providing that the original problem is convex) is available in [1].

The equivalent of original problem:

\[
\begin{align*}
\lambda \max_{x \geq 0, y \geq 0} & \quad x^2 + y^2 + \lambda (-x - y + 4) + \frac{\nu}{2} \| -x - y + 4 \|^2 \\
\lambda \min & \quad \text{subject to} \quad x \geq 0, y \geq 0
\end{align*}
\]
ADMM

The equivalent of original problem:

\[
\begin{align*}
\text{Maximize } & \quad \frac{1}{\lambda} \text{ Minimize } x^2 + y^2 + \lambda(-x - y + 4) + \frac{\gamma}{2} \| -x - y + 4 \|^2 \\
& \quad \text{subject to } x \geq 0, y \geq 0
\end{align*}
\]

Using ADMM, the problem above in iteration \( \nu \) can be decomposed to two subproblems:

\[
\begin{align*}
\text{Minimize } & \quad x^{(\nu)} - \lambda^{(\nu-1)} x^{(\nu)} + \frac{\gamma}{2} \| -x^{(\nu)} - y^{(\nu-1)} + 4 \|^2 \\
\text{Minimize } & \quad y^{(\nu)} - \lambda^{(\nu-1)} y^{(\nu)} + \frac{\gamma}{2} \| -y^{(\nu)} - x^{(\nu-1)} + 4 \|^2
\end{align*}
\]
ADMM

The equivalent of original problem:

$$\begin{align*}
\text{Maximize} & \quad \lambda \left( x^2 + y^2 + \lambda(-x - y + 4) + \frac{\gamma}{2} \| -x - y + 4 \|^2 \right) \\
\text{Minimize} & \quad \frac{\lambda}{x \geq 0, y \geq 0} \\
\end{align*}$$

Using ADMM, the problem above in iteration $v$ can be decomposed to two subproblems:

$$\begin{align*}
\text{Minimize} & \quad x^{2(v)} - \lambda^{(v-1)} x^{(v)} + \frac{\gamma}{2} \| -x^{(v)} - y^{(v-1)} + 4 \|^2 \\
\text{Minimize} & \quad y^{2(v)} - \lambda^{(v-1)} y^{(v)} + \frac{\gamma}{2} \| -y^{(v)} - x^{(v-1)} + 4 \|^2 \\
\end{align*}$$

In the first subproblem above (i.e., $x$-update), $x^{(v)}$ is variable, while $\lambda^{(v-1)}$ and $y^{(v-1)}$ are parameters.
**ADMM**

The equivalent of original problem:

\[
\begin{align*}
\max_{\lambda} \quad & \min_{x \geq 0, y \geq 0} x^2 + y^2 + \lambda (-x - y + 4) + \frac{\nu}{2} \| -x - y + 4 \|^2 \\
\end{align*}
\]

Using ADMM, the problem above in iteration \(\nu\) can be decomposed to two subproblems:

\[
\begin{align*}
\text{Minimize } x^{(v)} & - \lambda^{(v-1)} x^{(v)} + \frac{\nu}{2} \| -x^{(v)} - y^{(v-1)} + 4 \|^2 \\
\text{Minimize } y^{(v)} & - \lambda^{(v-1)} y^{(v)} + \frac{\nu}{2} \| -y^{(v)} - x^{(v-1)} + 4 \|^2 \\
\end{align*}
\]

In the second subproblem above (i.e., \(y\)-update), \(y^{(v)}\) is variable, while \(\lambda^{(v-1)}\) and \(x^{(v-1)}\) are parameters.
ADMM

The equivalent of original problem:

\[
\begin{align*}
\text{Maximize} & \quad \text{Minimize } x^2 + y^2 + \lambda(-x - y + 4) + \frac{\gamma}{2} \| -x - y + 4 \|^2 \\
\lambda & \quad x \geq 0, y \geq 0
\end{align*}
\]

Using ADMM, the problem above in iteration \( v \) can be decomposed to two subproblems:

\[
\begin{align*}
\text{Minimize } & \quad x^{2(v)} - \lambda^{(v-1)} x^{(v)} + \frac{\gamma}{2} \| -x^{(v)} - y^{(v-1)} + 4 \|^2 \\
& \quad x^{(v)} \geq 0 \\
\text{Minimize } & \quad y^{2(v)} - \lambda^{(v-1)} y^{(v)} + \frac{\gamma}{2} \| -y^{(v)} - x^{(v-1)} + 4 \|^2 \\
& \quad y^{(v)} \geq 0
\end{align*}
\]

\( \lambda \)-update: \( \lambda^{(v)} \leftarrow \lambda^{(v-1)} + \gamma(-x^{(v)} - y^{(v)} + 4) \)
ADMM

The equivalent of original problem:

\[
\begin{align*}
\text{Maximize} & \quad \lambda \frac{x^2}{x \geq 0, y \geq 0} + y^2 + \lambda (-x - y + 4) + \frac{\gamma}{2} \| -x - y + 4 \|^2 \\
\text{Minimize} & \quad \frac{x^2(v)}{x^{(v)} \geq 0} - \lambda^{(v-1)} x^{(v)} + \frac{\gamma}{2} \| -x^{(v)} - y^{(v-1)} + 4 \|^2 \\
\text{Minimize} & \quad \frac{y^2(v)}{y^{(v)} \geq 0} - \lambda^{(v-1)} y^{(v)} + \frac{\gamma}{2} \| -y^{(v)} - x^{(v-1)} + 4 \|^2 \\
\lambda \text{-update:} & \quad \lambda^{(v)} \leftarrow \lambda^{(v-1)} + \gamma \left( -x^{(v)} - y^{(v)} + 4 \right)
\end{align*}
\]

Algorithm: in each iteration, solve each subproblem and then update dual variable until convergence, i.e., when the primal residual (i.e., the value of penalty) is negligible, and therefore the value of dual variable does not change anymore.
Optional Assignment 2

Provide a GAMS (or Python or Julia) code for the previous illustrative example solved by ADMM algorithm!
Thanks for your attention!

Email: seykaz@elektro.dtu.dk