Primal and Dual Decomposition: Theory and Distributed Algorithms

Chee Wei Tan

Convex Optimization and its Applications to Computer Science

Outline

- Distributed Algorithms
- Primal Decomposition
- Lagrange Dual Decomposition

Distributed Algorithms

In previous lectures, we have studied optimization algorithms like

- Descent algorithm, e.g., Gradient method, Newton method
- Interior point algorithm
- Cutting plane algorithm
- Fixed-point iteration algorithm, e.g., power method

Distributed algorithms are preferred because:

- It's scalable
- It's robust
- Centralized command is not feasible or is too costly

Key issues:

- Local computation vs. global communication
- Scope, scale, and physical meaning of communication overhead
- Theoretical issues: Convergence? Optimality? Speed?
- Practical issues: Robustness? Synchronization? Complexity? Stability?
- Problem separability structure for decomposition: vertical and horizontal

Decomposition: LP Example

LP with variables u, v:

$$
\begin{array}{ll}
\text{maximize} & c_1^T u + c_2^T v \\
\text{subject to} & A_1 u \preceq b_1 \\
& A_2 v \preceq b_2 \\
& F_1 u + F_2 v \preceq h
\end{array}
$$

Coupling constraint: $F_1 u + F_2 v \preceq h$. Otherwise, separable into two LP

Primal Decomposition

Introduce variable z and rewrite coupling constraint as

$$F_1 u \preceq z, \quad F_2 v \preceq h - z$$

LP decomposed into a master problem and two subproblems:

$$\text{minimize}_z \phi_1(z) + \phi_2(z)$$

where

$$\begin{aligned} \phi_1(z) &= \inf_u \{c_1^T u | A_1 u \preceq b_1, Fu \preceq z\} \\ \phi_2(z) &= \inf_v \{c_2^T v | A_2 v \preceq b_2, F_2 v \preceq h - z\} \end{aligned}$$

Subgradient of function $f : \mathbf{R}^n \to \mathbf{R}$ at x is a vector g such that

$$f(y) \geq f(x) + g^T (y - x), \ \ \forall y$$

Primal Decomposition

For each iteration t:

1. Solve two separate LPs to obtain optimal $u(t), v(t)$ and associated dual variables $\lambda_1(t), \lambda_2(t)$
2. Subgradient update: $g(t) = -\lambda_1(t) + \lambda_2(t)$
3. Master algorithm update: $z(t+1) = z(t) - \alpha(t)g(t)$ where $\alpha(t) \geq 0,\ \lim_{t\to\infty} \alpha_t = 0$ and $\sum_{t=1}^{\infty} \alpha(t) = \infty$

Interpretation:

- z fixes allocation of resources between two subproblems and master problem iteratively finds best allocation of resources
- More of each resource is allocated to the subproblem with larger Lagrange dual variable at each step

Dual Decomposition

Form partial Lagrangian:

$$
\begin{aligned}
L(u, v, \lambda) &= c_1^T u + c_2^T v + \lambda^T (F_1 u + F_2 v - h) \\
&= (F_1^T \lambda + c_1)^T u + (F_2^T \lambda + c_2)^T v - \lambda^T h
\end{aligned}
$$

Dual function:

$$
\begin{aligned}
q(\lambda) &= \inf_{u,v} \{ L(u, v, \lambda) | A_1 u \preceq b1, A_2 v \preceq b_2 \} \\
&= -\lambda^T h + \inf_{u: A_1 u \preceq b_1} (F_1^T \lambda + c_1)^T u + \inf_{v: A_2 v \preceq b_2} (F_2^T \lambda + c_2)^T v
\end{aligned}
$$

Dual problem:

$$
\begin{array}{ll}
\text{maximize} & q(\lambda) \\
\text{subject to} & \lambda \succeq 0
\end{array}
$$

Dual Decomposition

Solve the following LP in u, with minimizer $u^*(\lambda(t))$

$$\begin{array}{ll} \text{minimize} & (F_1^T\lambda(t)+c_1)^T u \\ \text{subject to} & A_1 u \preceq b_1 \end{array}$$

Solve the following LP in v, with minimizer $v^*(\lambda(t))$

$$\begin{array}{ll} \text{minimize} & (F_2^T\lambda(t)+c_2)^T v \\ \text{subject to} & A_2 v \preceq b_2 \end{array}$$

Use the following subgradient (to $-q$) to update λ:

$$g(t) = -F_1 u^*(\lambda(t)) - F_2 v^*(\lambda(t)) + h, \quad \lambda(t+1) = [\lambda(t) - \alpha(t) g(t)]^+$$

Interpretation:

Master algorithm adjusts prices λ, which regulates the separate solutions of two subproblems

Summary

- Decouple a coupling constraint: primal or dual decomposition

- Dual decomposition algorithm: a large-scale engineering application in the next lecture

Reading assignment:

- Sections 3.4, 7.5 in D. P. Bertsekas and J. N. Tsitsiklis, Parallel and Distributed Computation: Numerical Methods, Athena Scientific 1999.