Price Directed Distributed Optimization and Primal Recovery

Tetiana Parshakova Yicheng Bai Garrett van Ryzin Stephen Boyd

Amazon SCOT

INFORMS

10/21/24

Outline

Our focus

our setting

- substantial computational cost per agent
- small number of iterations is desirable
- care about feasibility, some suboptimality is tolerable

we propose (postprocessing) method that

- uses low-precision optimal dual variable
- recovers close to feasible primal point
- uses only parallel calls to agents; avoids sequential calls

Distributed optimization

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & Ax \leq b \end{array}$$

- $x = (x_1, \ldots, x_M) \in \mathbf{R}^n$ is variable, $x_i \in \mathbf{R}^{n_i}$
- $f(x) = \sum_{i=1}^M f_i(x_i)$ is block separable
- $f_i : \mathbf{R}^{n_i} \to \mathbf{R} \cup \{\infty\}$ convex, closed and proper
- infinite values of f encode constraints
- $A = (A_1, \ldots, A_M) \in \mathbf{R}^{m \times n}$ and $b \in \mathbf{R}^m$ are given

Conjugate subgradient oracle

- for $y \in \mathbf{R}^m$ ($\mathbf{dom}\, f^* = \mathbf{R}^m$), oracle returns $x(y) \in \partial f^*(y)$

$$-f^*(y) = \inf_{x \in \mathbf{dom}\, f} \left(f(x) - y^T x\right)$$

- no access to function values $f(x)$ or subgradient in $\partial f(x)$
- neutral cutting plane for dual variable

$$\left\{\tilde{\lambda} \in \mathbf{R}^m \mid (-Ax(y) + b)^T(\tilde{\lambda} - \lambda) \le 0,\ y = -A^T\lambda\right\}$$

Dual problem

solve the dual problem

- subgradient methods (Shor, 1962)
- localization methods
  - analytic center cutting-plane method (ACCPM)
  - maximum volume ellipsoid cutting-plane method
  - Chebyshev center cutting-plane method
- we've settled on proximal point ACCPM

Optimality conditions

- KKT conditions

$$x \in \partial f^*(y),\ \ y = -A^T\lambda \tag{1}$$
$$Ax \le b \tag{2}$$
$$\lambda_j (Ax - b)_j = 0, \quad j = 1, \ldots, m \tag{3}$$
$$\lambda \ge 0 \tag{4}$$

- primal and complementary slackness residuals for a pair $(x, \lambda) \in \mathbf{R}^n \times \mathbf{R}^m_+$

$$r_p = \mathbf{1}^T (Ax - b)_+, \qquad r_c = \lambda^T |Ax - b|$$

Outline

Primal recovery

given current prices $y^k = -A^T\lambda^k = (y_1^k, \ldots, y_M^k)$

1. agent i returns multiple offers $z_i^{(1)}, \ldots, z_i^{(N)}$ wrt $y_i^k = -A_i^T\lambda^k$
   - $-f_i^*(y_i^k) \le f_i(z_i) - y_i^{kT} z_i \approx -f_i^*(y_i^k)$
   - computed in parallel (clock time equivalent to single response)

2. central node constructs a convex combination of offers $\bar{x}$ minimizing the residuals

$$
\begin{array}{ll}
\text{minimize} & r_p + r_c \\
\text{subject to} & \bar{x}_i = Z_i u_i, \quad i = 1, \ldots, M \\
& \mathbf{1}^T u_i = 1, \; u_i \ge 0, \quad i = 1, \ldots, M \\
& \bar{x} = (\bar{x}_1, \ldots, \bar{x}_M)
\end{array}
$$

don't need to run primal recovery at every step k (periodic runs sufficient)

Approximate conjugate subgradient oracle

- for any y, an ϵ_v-value suboptimal primal variable $x^{\mathsf{v}}(y)$

$$-f^*(y) \leq f(x^{\mathsf{v}}(y)) - y^T x^{\mathsf{v}}(y) \leq -f^*(y) + \epsilon_v |f^*(y)|$$

  - any convex combination $\bar{x}$: $f(\bar{x}) - y^T \bar{x} \leq -f^*(y) + \epsilon_v |f^*(y)|$

- for any y, primal variable $x^{\mathsf{p}}(y)$ with ϵ_p-perturbed prices

$$f(x^{\mathsf{p}}(y)) - (y+\delta)^T x^{\mathsf{p}}(y) = -f^*(y+\delta), \quad \delta \in [-\epsilon_p |y|, \epsilon_p |y|]$$

  - for L-Lipschitz f^*, for any convex combination $\bar{x}$: $f(\bar{x}) - y^T \bar{x} \leq -f^*(y) + \epsilon_p L \|y\|$

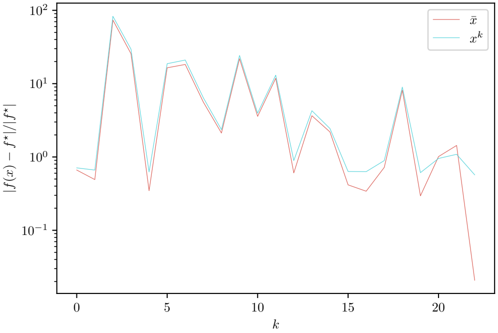
Outline

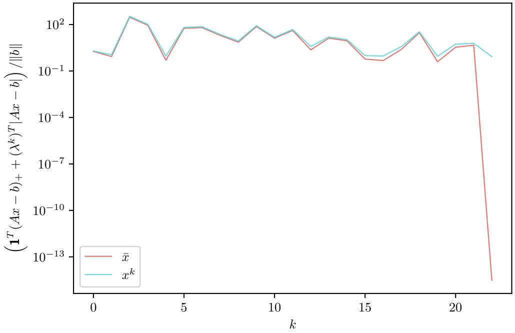
Setting

- run localization method for solving the dual, compute λ^k
- price directed interface gives $x^k \in \partial f^*(-A^T\lambda^k)$
- $\bar{x}$ returned by primal recovery method
- $N = 10$ suboptimal offers with $\epsilon = 10\%$-suboptimality
- primal recovery effective when $m \ll n$

Resource allocation

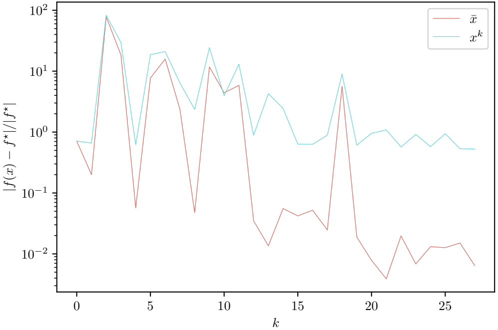
- $M = 100$ agents, $m = 8$ resources, $n = 800$
- $\epsilon_v = 10\%$-value suboptimality

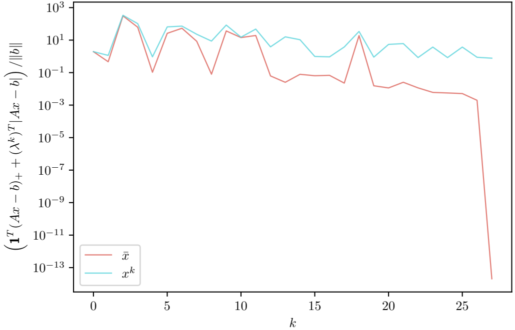




Resource allocation

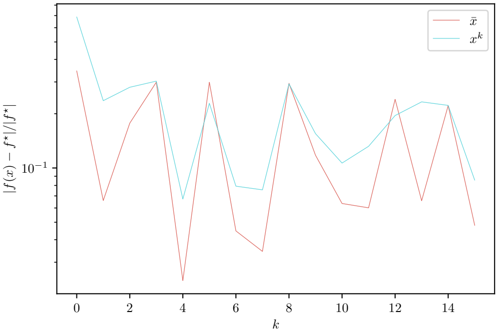
- $M = 100$ agents, $m = 8$ resources, $n = 800$
- $\epsilon_p = 10\%$-perturbed prices

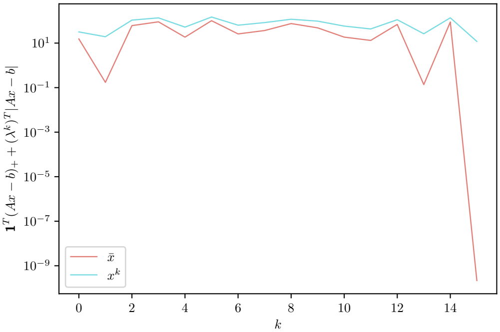




Assignment problem

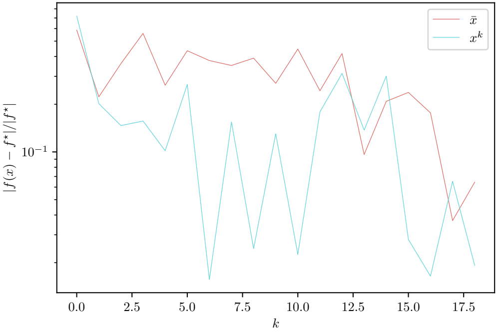
- $M = 208$ agents, $m = 8$, $n = 1608$
- $\epsilon_v = 10\%$-value suboptimality

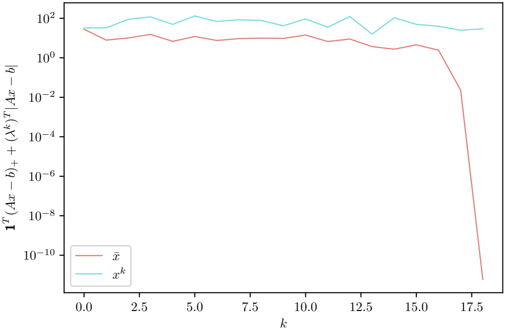




Assignment problem

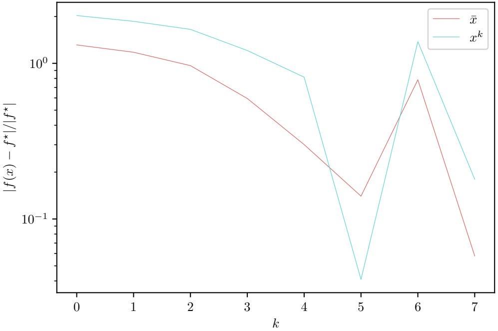
- $M = 208$ agents, $m = 8$, $n = 1608$
- $\epsilon_p = 10\%$-perturbed prices

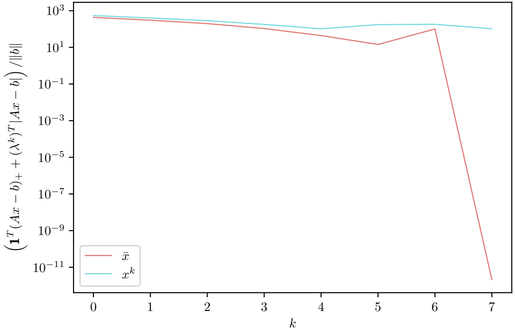




Linear programming

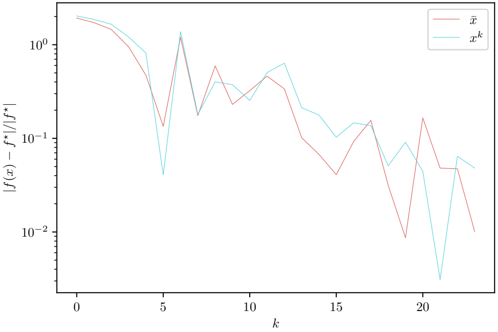
- $M = 100$ agents, $m = 8$, $n = 800$
- $\epsilon_v = 10\%$-value suboptimality

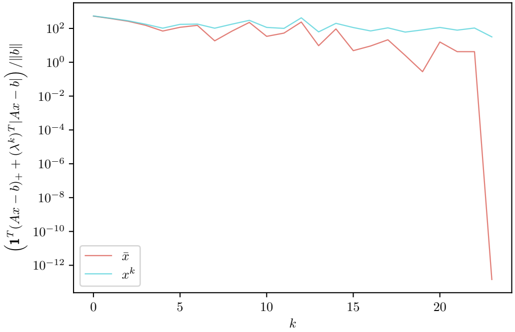




Linear programming

- $M = 100$ agents, $m = 8$, $n = 800$
- $\epsilon_p = 10\%$-perturbed prices





Outline

Primal recovery method

- lower bound

$$\mathcal{L}(x^k, \lambda^k) = \underset{x}{\operatorname{argmin}}(f(x) + \lambda^{kT}(Ax - b)) = -f^*(-A^T\lambda^k) - \lambda^{kT}b \leq p^\star$$

- ϵ_v-value suboptimality

$$\mathcal{L}(\bar{x}^{\mathsf{v}}, \lambda^k) = f(\bar{x}^{\mathsf{v}}) + \lambda^{kT}(A\bar{x}^{\mathsf{v}} - b) \leq p^\star + \epsilon_v \sum_{i=1}^{M} |f_i^*(-A_i^T\lambda^k)|$$

- ϵ_p-price perturbation

$$\mathcal{L}(\bar{x}^{\mathsf{p}}, \lambda^k) = f(\bar{x}^{\mathsf{p}}) + \lambda^{kT}(A\bar{x}^{\mathsf{p}} - b) \leq p^\star + \epsilon_p L \sum_{i=1}^{M} \|A_i^T\lambda^k\|$$