Simple/multi-level fast direct solver for IE

Ondrej Maxian, Anqi Mao

October 28, 2020

Outline

1. CH13: A simple direct solver for integral equations
   - Introduction
   - Block separable matrices

Background

Formulate a physical problem in the form of an integral equation such as

$$\alpha\, q(\boldsymbol{x}) + \int_\Gamma k(\boldsymbol{x},\boldsymbol{y})\, q(\boldsymbol{y})\, ds(\boldsymbol{y}) = f(\boldsymbol{x}), \qquad \boldsymbol{x} \in \Gamma,$$

and then discretize it to obtain a linear system

$$\mathbf{Aq} = \mathbf{f} \tag{1}$$

Question: how to solve the linear system efficiently?

Solvers for discretized IE

Usually, the coefficient matrix **A** is dense.

Gaussian elimination: cubic complexity

iterative solver: $T_{solve} \simeq N_{iter} \times T_{matvec}$

T_{matvec}: linear complexity (FMM, $\mathcal{H}$-matrix)

N_{iter}: preconditioner

fast direct solvers: linear complexity

build stage: build approximate inverse **B** of **A**

solve stage: compute approximate solution $\mathbf{q}_{approx} = \mathbf{Bf}$

compelling in solving a sequence of linear systems

simple single-level scheme $\rightarrow$ multi-level scheme

Block separable matrices

$\mathbf{A} : N \times N$. Tessellate it into $p \times p$ blocks. Each block: $n \times n$.

$$\mathbf{A} = \begin{bmatrix} \mathbf{D}_1 & \mathbf{A}_{1,2} & \mathbf{A}_{1,3} & \cdots & \mathbf{A}_{1,p} \\ \mathbf{A}_{2,1} & \mathbf{D}_2 & \mathbf{A}_{2,3} & \cdots & \mathbf{A}_{2,p} \\ \vdots & \vdots & \vdots & & \vdots \\ \mathbf{A}_{p,1} & \mathbf{A}_{p,2} & \mathbf{A}_{p,3} & \cdots & \mathbf{D}_p \end{bmatrix}$$

Assume

- upper bound k for ranks of all off-diagonal blocks
- basis matrices $\{\mathbf{U}_k\}_{k=1}^{p}$ and $\{\mathbf{V}_k\}_{k=1}^{p}$ such that

$$\underset{n \times n}{\mathbf{A}_{\sigma,\tau}} = \underset{n \times k}{\mathbf{U}_\sigma} \; \underset{k \times k}{\tilde{\mathbf{A}}_{\sigma,\tau}} \; \underset{k \times n}{\mathbf{V}_\tau^*}, \qquad \sigma, \tau \in \{1, 2, \ldots, p\}, \quad \sigma \neq \tau.$$

Block separable matrices

Therefore, **A** admits a block factorization:

$$\underset{pn \times pn}{\mathbf{A}} = \underset{pn \times pk}{\mathbf{U}} \; \underset{pk \times pk}{\tilde{\mathbf{A}}} \; \underset{pk \times pn}{\mathbf{V}^*} + \underset{pn \times pn}{\mathbf{D}},$$

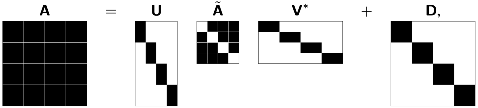
where

$$\mathbf{U} = \mathrm{diag}(\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_p),$$
$$\mathbf{V} = \mathrm{diag}(\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_p),$$
$$\mathbf{D} = \mathrm{diag}(\mathbf{D}_1, \mathbf{D}_2, \ldots, \mathbf{D}_p),$$

and

$$\tilde{\mathbf{A}} = \begin{bmatrix} \mathbf{0} & \tilde{\mathbf{A}}_{12} & \tilde{\mathbf{A}}_{13} & \cdots \\ \tilde{\mathbf{A}}_{21} & \mathbf{0} & \tilde{\mathbf{A}}_{23} & \cdots \\ \tilde{\mathbf{A}}_{31} & \tilde{\mathbf{A}}_{32} & \mathbf{0} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}.$$

For $p = 4$,



Variation of the Woodbury formula

Lemma (variation of the Woodbury formula)

Suppose that $\mathbf{A}$ is an invertible $N \times N$ matrix, K is a positive integer smaller than N, and $\mathbf{A}$ admits the factorization:

$$\underset{N \times N}{\mathbf{A}} = \underset{N \times K}{\mathbf{U}} \; \underset{K \times K}{\tilde{\mathbf{A}}} \; \underset{K \times N}{\mathbf{V}^*} + \underset{N \times N}{\mathbf{D}}.$$

Then

$$\underset{N \times N}{\mathbf{A}^{-1}} = \underset{N \times K}{\mathbf{E}} \; \underset{K \times K}{(\tilde{\mathbf{A}} + \hat{\mathbf{D}})^{-1}} \; \underset{K \times N}{\mathbf{F}^*} + \underset{N \times N}{\mathbf{G}},$$

where

$$\hat{\mathbf{D}} = (\mathbf{V}^* \mathbf{D}^{-1} \mathbf{U})^{-1},$$
$$\mathbf{E} = \mathbf{D}^{-1} \mathbf{U} \hat{\mathbf{D}},$$
$$\mathbf{F} = (\hat{\mathbf{D}} \mathbf{V}^* \mathbf{D}^{-1})^*,$$
$$\mathbf{G} = \mathbf{D}^{-1} - \mathbf{D}^{-1} \mathbf{U} \hat{\mathbf{D}} \mathbf{V}^* \mathbf{D}^{-1},$$

provided all inverses that appear exist. Moreover, $\text{rank}(\mathbf{G}) = N - K$.

Proof

Fix an $\mathbf{f}$. Set $\mathbf{q} = \mathbf{A}^{-1}\mathbf{f}$. Then $\mathbf{A}\mathbf{q} = \mathbf{f}$. Set $\tilde{\mathbf{q}} = \mathbf{V}^*\mathbf{q}$. We get a linear system

$$\begin{bmatrix} \mathbf{D} & \mathbf{U}\tilde{\mathbf{A}} \\ -\mathbf{V}^* & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{q} \\ \tilde{\mathbf{q}} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{0} \end{bmatrix}$$

From the first row, $\mathbf{q} = \mathbf{D}^{-1}\mathbf{f} - \mathbf{D}^{-1}\mathbf{U}\tilde{\mathbf{A}}\tilde{\mathbf{q}}$. Substituting into the second row yields

$$\left(\mathbf{I} + \mathbf{V}^*\mathbf{D}^{-1}\mathbf{U}\tilde{\mathbf{A}}\right)\tilde{\mathbf{q}} = \mathbf{V}^*\mathbf{D}^{-1}\mathbf{f}$$

Multiply both sides by $\hat{\mathbf{D}} = (\mathbf{V}^*\mathbf{D}^{-1}\mathbf{U})^{-1}$ to get

$$\left(\hat{\mathbf{D}} + \tilde{\mathbf{A}}\right)\tilde{\mathbf{q}} = \hat{\mathbf{D}}\mathbf{V}^*\mathbf{D}^{-1}\mathbf{f}.$$

Therefore, we can express $\mathbf{q}$ as

$$\begin{aligned}
\mathbf{q} &= \mathbf{D}^{-1}\mathbf{f} - \mathbf{D}^{-1}\mathbf{U}\tilde{\mathbf{A}}\tilde{\mathbf{q}} \\
&= \mathbf{D}^{-1}\mathbf{f} - \mathbf{D}^{-1}\mathbf{U}\left(\hat{\mathbf{D}}\mathbf{V}^*\mathbf{D}^{-1}\mathbf{f} - \hat{\mathbf{D}}\tilde{\mathbf{q}}\right) \\
&= \underbrace{\left(\mathbf{D}^{-1} - \mathbf{D}^{-1}\mathbf{U}\hat{\mathbf{D}}\mathbf{V}^*\mathbf{D}^{-1}\right)}_{=\mathbf{G}}\mathbf{f} + \underbrace{\mathbf{D}^{-1}\mathbf{U}\hat{\mathbf{D}}}_{=\mathbf{E}}\left(\hat{\mathbf{D}} + \tilde{\mathbf{A}}\right)^{-1}\underbrace{\hat{\mathbf{D}}\mathbf{V}^*\mathbf{D}^{-1}}_{=\mathbf{F}^*}\mathbf{f}
\end{aligned}$$

Proof: rank(G)$= N - K$

Observe that

$$\mathbf{V}^*\mathbf{G} = \mathbf{V}^*\mathbf{D}^{-1} - \mathbf{V}^*\mathbf{D}^{-1}\mathbf{U}\hat{\mathbf{D}}\mathbf{V}^*\mathbf{D}^{-1} = \mathbf{V}^*\mathbf{D}^{-1} - \hat{\mathbf{D}}^{-1}\hat{\mathbf{D}}\mathbf{V}^*\mathbf{D}^{-1}$$
$$= \mathbf{V}^*\mathbf{D}^{-1} - \mathbf{V}^*\mathbf{D}^{-1} = \mathbf{0}$$

Thus,

$$\text{rank}(\mathbf{V}^*) + \text{rank}(\mathbf{G}) - N \leq 0 \Rightarrow \text{rank}(\mathbf{G}) \leq N - K$$

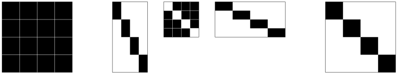
Also,

$$\text{rank}(\mathbf{A}^{-1}) \leq \text{rank}(\mathbf{E}(\tilde{\mathbf{A}} + \hat{\mathbf{D}})^{-1}\mathbf{F}^*) + \text{rank}(\mathbf{G}) \Rightarrow \text{rank}(\mathbf{G}) \geq N - K$$

Apply the lemma

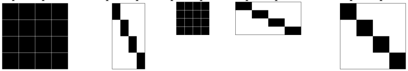
Recall the block structure of $\mathbf{A}$, for $p = 4$

$$\mathbf{A} = \mathbf{U} \tilde{\mathbf{A}} \mathbf{V}^* + \mathbf{D},$$



Compute $\mathbf{A}^{-1}$ by the lemma

$$\underset{pn \times pn}{\mathbf{A}^{-1}} = \underset{pn \times pk}{\mathbf{E}} \underset{pk \times pk}{(\tilde{\mathbf{A}} + \hat{\mathbf{D}})^{-1}} \underset{pk \times pn}{\mathbf{F}^*} + \underset{pn \times pn}{\mathbf{G}}.$$



$$\hat{\mathbf{D}} = \left(\mathbf{V}^* \mathbf{D}^{-1} \mathbf{U}\right)^{-1},$$
$$\mathbf{E} = \mathbf{D}^{-1} \mathbf{U} \hat{\mathbf{D}},$$
$$\mathbf{F} = (\hat{\mathbf{D}} \mathbf{V}^* \mathbf{D}^{-1})^*,$$
$$\mathbf{G} = \mathbf{D}^{-1} - \mathbf{D}^{-1} \mathbf{U} \hat{\mathbf{D}} \mathbf{V}^* \mathbf{D}^{-1}$$

- $\hat{\mathbf{D}}, \mathbf{E}, \mathbf{F}, \mathbf{G}$ are cheap to form (block diagonal)
- invert $pn \times pn$ matrix $\mathbf{A} \Rightarrow$ invert small $pk \times pk$ matrix $\tilde{\mathbf{A}} + \hat{\mathbf{D}}$

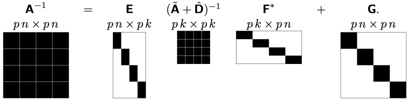
Remark

A more common version of the Woodbury formula

$$\left(\mathbf{D} + \mathbf{U}\tilde{\mathbf{A}}\mathbf{V}^*\right)^{-1} = \mathbf{D}^{-1} - \mathbf{D}^{-1}\mathbf{U}\left(\tilde{\mathbf{A}}^{-1} + \mathbf{V}^*\mathbf{D}^{-1}\mathbf{U}\right)^{-1}\mathbf{V}^*\mathbf{D}^{-1}.$$

- If both $\tilde{\mathbf{A}}$ and $\mathbf{V}^*\mathbf{D}^{-1}\mathbf{U}$ are invertible, then the two versions are equivalent
- $\tilde{\mathbf{A}}$ is not block diagonal, often not invertible

Asymptotic complexity



Suppose we are given the factors in a block separable factorization.

- compute $\hat{\mathbf{D}}, \mathbf{E}, \mathbf{F}, \mathbf{G}$: pn^3
- invert $\tilde{\mathbf{A}} + \hat{\mathbf{D}}$: $(pk)^3$

Suppose that k is fixed. What is the optimal choice of p?

$$T \sim p\,(N/p)^3 + (pk)^3 \sim p^{-2}N^3 + p^3k^3$$

Optimal choice $p \sim (N/k)^{3/5}$, which leads to

$$T \sim N^3p^{-2} + k^3p^3 \sim N^3(N/k)^{-6/5} + k^3(N/k)^{9/5} \sim N^{9/5}k^{6/5}$$

Compute a block separable representation

Definition

Let $\mathbf{A}$ be an $N \times N$ matrix, let $I = 1 : N$ be its index vector, and let $I = I_1 \cup I_2 \cup \ldots \cup I_p$ be a disjoint partition of I. Then $\mathbf{A}$ has block rank k w.r.t this partition if there exist matrices $\{\mathbf{U}_\tau\}_{\tau=1}^{p}$ and $\{\mathbf{V}_\tau\}_{\tau=1}^{p}$ such that each off-diagonal block of $\mathbf{A}$ admits a factorization

$$\underset{n_\sigma \times n_\tau}{\mathbf{A}(I_\sigma, I_\tau)} = \underset{n_\sigma \times k}{\mathbf{U}_\sigma} \; \underset{k \times k}{\tilde{\mathbf{A}}_{\sigma,\tau}} \; \underset{k \times n_\tau}{\mathbf{V}_\tau^*}, \qquad \sigma, \tau \in \{1, 2, \ldots, p\}, \quad \sigma \neq \tau,$$

where n_τ is the length of I_τ.

block separable: block rank k is small

block separable (to precision ϵ): block rank k is small (within preset tolerance ϵ)

Compute a block separable representation

For any given τ, the columns of $\mathbf{U}_\tau$ span the columns of every block $\mathbf{A}_{\tau,\sigma}$ for $\sigma \neq \tau$.
In other words, the columns of $\mathbf{U}_\tau$ span all the columns in the matrix

$$\mathbf{A}(I_\tau, I_\tau^{\mathrm{c}}) = \left[\mathbf{A}_{\tau,1},\, \mathbf{A}_{\tau,2},\, \dots,\, \mathbf{A}_{\tau,\tau-1},\, \mathbf{A}_{\tau,\tau+1},\, \dots,\, \mathbf{A}_{\tau,p}\right]$$

where $I_\tau^c = I \backslash I_\tau$.
Solve the task using SVD.
When a block separable matrix arises from the discretization of a boundary integral operator, there exist compression techniques of linear complexity.