Numerical Comparison Summary

Qiyuan Pang

July 19, 2021

The following quantities are used in the rest of the section to evaluate the performance of the preconditioner:

- N: problem size;
- e_a: the relative error set for the butterfly approximation $\hat{K}$ of K;
- ϵ: the fixed tolerance set in HIF/HQR;
- e_f: forward error of a factorization (HIF/HQR) (e.g., $\hat{A}$ factorizes A, then $e_f = \|\hat{A}x - Ax\|/\|Ax\|$);
- e_h: the accuracy of HODLR construction using the peeling algorithm.
- e_s: the relative error of the approximation $\hat{G}\hat{K}^*$ of K^{-1}, defined as $\|\hat{G}\hat{K}^*b - x\|/\|x\|$ where x is a random vector and $b = Kx$;
- n_i: the number of iterations used in PCG until covergence;
- e: the relative error of the solution returned by PCG.

Among all experiments below, the stopping criteria set for PCG is tolerance $1e-10$.

Examples (1D). We begin with an example of 1D discrete FIO of the form

$$u(x) = \int_{\mathbb{R}} a(x)e^{2\pi i\Phi(x,\xi)}\hat{f}(\xi)d\xi$$

There are five 1D kernels to test here, as follows:

$$a = 1, \Phi(x,\xi) = x\cdot\xi + c(x)|\xi|, c(x) = (2+\sin(2\pi x))/8, \tag{1}$$

$$a = 1, \Phi(x,\xi) = x\cdot\xi + c(x)\xi, c(x) = (2+\sin(2\pi x))/5.94, \tag{2}$$

$$a = \sum_{k=0}^{n_k} e^{-\frac{(x-x_k)^2+(\xi-\xi_k)^2}{\sigma^2}}, \sigma = 0.05, \Phi(x,\xi) = x\cdot\xi + c(x)|\xi|, c(x) = (2+\sin(2\pi x))/8, \tag{3}$$

$$a = \sum_{k=0}^{n_k} e^{-\frac{(x-x_k)^2+(\xi-\xi_k)^2}{\sigma^2}}, \sigma = 0.1, \Phi(x,\xi) = x\cdot\xi + c(x)|\xi|, c(x) = (2+\sin(2\pi x))/8, \tag{4}$$

$$a = \sum_{k=0}^{n_k} e^{-\frac{(x-x_k)^2+(\xi-\xi_k)^2}{\sigma^2}}, \sigma = 0.04, \Phi(x,\xi) = x\cdot\xi + c(x)|\xi|, c(x) = (2+\sin(2\pi x))/7, \tag{5}$$

Note that the amplitude function a in (3), (4), and (5) are as the same as that in Example 2 in Lexing's preprint. Here we skip the exact formula of a.

Discretizing x and ξ on $[0,1)$ and $[-N/2, N/2)$ with N points,

$$x_i = (i-1)/N, \xi_j = j - 1 - N/2.$$

leads to the discrete system $u = Kf$.

Regularization. Here we consider the following L1 optimization problem:

$$\min_f |\Phi(f)| + \|Kf - u\|_2^2,$$

which is equal to

$$\min_{f,d} |d| + \|Kf-u\|_2^2 \quad such \quad that \quad d = \Phi(f).$$

To solve this, first convert it into an unconstrained problem:

$$\min_{f,d} |d| + \|Kf-u\|_2^2 + \frac{\lambda}{2}\|d - \Phi(f)\|_2^2.$$

This could be solved using the Split Bregman Iteration:

$$\begin{aligned} Step1 : f^{k+1} &= \min_f \|Kf-u\|_2^2 + \frac{\lambda}{2}\|d^k - \Phi(f) - b^k\|_2^2 \\ Step2 : d^{k+1} &= \min_d |d| + \frac{\lambda}{2}\|d - \Phi(f^{k+1}) - b^k\|_2^2 \\ Step3 : b^{k+1} &= b^k + \Phi(f^{k+1}) - d^{k+1} \end{aligned} \tag{6}$$

The iteration stops when $\|f^{k+1} - f^k\| < tol$, and $tol = 1E-10$ in the experiments. $\lambda = 1E-8$. And the real f is set to be sparse in the experiments.

In Step 2, we can explicitly compute the optimal value of d using shrinkage operators. We simply compute

$$d_j^{k+1} = shrink(\Phi(f)_j + b_j^k, 1/\lambda), \tag{7}$$

where

$$shrink(x,\gamma) = \frac{x}{|x|} \times \max(|x| - \gamma, 0). \tag{8}$$

In Step 1, PCG or Gauss-Seidel (GS) method could be used.

L1 norm: $\Phi(f) = f$. Then the regularizer is $|f|$, and Step 1 becomes:

$$(2K^*K + \lambda)f^{k+1} = 2K^*u + \lambda(d^k - b^k) \tag{9}$$

In the experiments, this is solved using PCG with an approximation of $(2K^*K+\lambda)^{-1}$ as a preconditioner. And an approximation of $(K^*K)^{-1}K^*u$ computed using the HIF idea in Ying's preprint is used to initialize both f^0 and d^0. b^0 is initialized as a zero vector.

TV norm: $\Phi(f) = \nabla f$. Since f is from the frequency domain (like $[-N/2, N/2]$), then it is not possible to compute its derivatives using finite difference method. To overcome this, we could tranform the frequency domain to the space domain (like $[0,1]$) using FFT, i.e., $f = Fg$ where F is the discrete Fourier tranform. Now we should solve $KFg = u$. Then the optimization problem becomes:

$$\min_{g,d} |d| + \|KFg - u\|_2^2 + \frac{\lambda}{2}\|d - \nabla g\|_2^2,$$

and Step 1 becomes:

$$(2F^*K^*KF - \lambda\Delta)g^{k+1} = 2F^*K^*u + \lambda\nabla(d^k - b^k). \tag{10}$$

In the experiments, this is solved using PCG with $(2F^*K^*KF)^{-1}$ as a preconditioner and g^k as an initial guess. We could also compute $L+U = 2F^*K^*KF - \lambda\Delta$ using HODLR representation such that L^{-1} and U could be applied quickly. Then Gauss-Seidel method could be applied to solve the equation for g^{k+1}. And an approximation of $(F^*K^*KF)^{-1}K^*F^*u$ computed using the HIF idea in Ying's preprint is used to initialize both g^0 and d^0. b^0 is initialized as a zero vector.

Table 2 summarizes the results for 1D kernel (1). Table 3 summarizes the results for 1D kernel (2). Table 4 summarizes the results for 1D kernel (5).

N	cond	Kernel 1	Kernel 2	Kernel 3	Kernel 4	Kernel 5
2^8	A	1.0660e+02	3.4571e+02	5.8246e+02	2.6794e+02	1.6771e+03
	A^*A	1.1364e+04	1.1952e+05	3.3926e+05	7.1790e+04	2.8128e+06
2^9	A	1.1372e+02	6.7118e+02	3.7644e+02	1.6732e+02	3.0517e+03
	A^*A	1.2932e+04	4.5048e+05	1.4171e+05	2.7995e+04	9.3131e+06
2^{10}	A	1.0870e+02	3.9350e+03	4.4032e+02	1.8616e+02	3.3981e+03
	A^*A	1.1815e+04	1.5484e+07	1.9388e+05	3.4657e+04	1.1547e+07
2^{11}	A	1.1999e+02	5.2064e+05	4.3108e+02	1.9745e+02	5.0709e+03
	A^*A	1.4398e+04	2.7107e+11	1.8583e+05	3.8988e+04	2.5714e+07
2^{12}	A	1.2626e+02	1.3755e+10	5.5073e+02	1.9459e+02	3.4614e+03
	A^*A	1.5943e+04	2.9863e+17	3.0330e+05	3.7866e+04	1.1981e+07

Table 1: Condition numbers of all kernels

	$\hat{K} \approx K$	HIF				$L1$		$TV(PCG)$		$TV(GS)$	
N	e_a	e_h	ϵ	e_f	e_s	n_i	e	n_i	e	n_i	e
2^8	9.4e-16	1.3e-08	1e-7	1.0e-08	1.5e-08	2	1.4e-08	80	3.5e-07	80	3.5e-07
2^9	1.4e-15	1.0e-08	1e-7	1.1e-08	1.1e-08	2	1.6e-08	80	3.6e-07	80	3.6e-07
2^{10}	1.5e-15	1.0e-08	1e-7	7.7e-09	1.5e-08	2	1.0e-08	80	3.4e-07	80	3.4e-07
2^{11}	2.4e-10	9.4e-09	1e-7	9.2e-09	1.1e-08	2	1.3e-08	80	3.3e-07	80	3.3e-07
2^{12}	3.6e-10	9.8e-09	1e-7	7.7e-09	1.2e-08	2	1.0e-08	80	2.8e-07	80	2.8e-07
2^{13}	1.9e-10	8.7e-09	1e-7	1.2e-08	9.8e-09	2	1.6e-08	80	3.5e-07	80	3.5e-07
2^{14}	1.9e-10	9.1e-09	1e-7	8.4e-09	1.2e-08	2	1.1e-08	80	3.3e-07	80	3.0e-07
2^{15}	5.1e-10	9.2e-09	1e-7	8.2e-09	1.2e-08	2	1.1e-08	80	3.0e-07	80	3.8e-07
2^{16}	4.6e-10	9.2e-09	1e-7	9.5e-09	1.1e-08	2	1.3e-08	80	3.0e-07	80	3.5e-07

Table 2: Numerical comparison between different regularizers for kernel (1). $\lambda = 1E-8$. The iteration stops when $\|f^{k+1} - f^k\| < tol = 1E-10$.

	$\hat{K} \approx K$	HIF				$L1$		$TV(PCG)$		$TV(GS)$	
N	e_a	e_h	ϵ	e_f	e_s	n_i	e	n_i	e	n_i	e
2^8	9.1e-16	2.8e-08	1e-7	1.8e-08	1.1e-07	2	9.7e-08	80	4.5e-07	80	4.5e-07
2^9	1.5e-15	1.2e-07	1e-7	1.1e-07	2.9e-07	2	8.2e-07	80	9.0e-07	80	7.6e-07
2^{10}	1.6e-15	4.8e-08	1e-7	1.1e-07	2.1e-05	2	2.9e-04	80	3.1e-04	80	1.5e-04
2^{11}	4.4e-10	6.3e-08	1e-7	1.2e-07	5.5e-02	1	2.7e+00	1	8.3e+00	80	8.3e+00
2^{12}	4.0e-10	9.6e-08	1e-7	1.1e-07	4.1e-02	1	1.6e+01	1	7.6e-01	80	7.6e-01
2^{13}	1.8e-10	2.0e-05	1e-7	1.4e-05	5.9e-02	2	8.9e-01	1	9.1e-01	80	1.9e+00
2^{14}	2.2e-10	2.8e-01	1e-7	1.9e-01	4.7e+00	2	1.3e+01	1	7.2e+00	80	NaN
2^{15}	5.5e-10	6.3e-01	1e-7	5.4e-01	5.9e+01	1	7.9e+01	1	2.0e+01	80	NaN
2^{16}	5.0e-10	6.5e-01	1e-7	4.3e-01	1.9e+01	1	1.0e+02	80	4.9e+01	80	NaN

Table 3: Numerical comparison between different regularizers for kernel (2). $\lambda = 1E-8$. The iteration stops when $\|f^{k+1} - f^k\| < tol = 1E-10$.

	$\hat{K} \approx K$	HIF				$L1$		$TV(PCG)$		$TV(GS)$	
N	e_a	e_h	ϵ	e_f	e_s	n_i	e	n_i	e	n_i	e
2^8	9.0e-16	1.3e-08	1e-7	1.8e-08	7.8e-06	2	1.2e-04	80	4.9e-03	80	1.1e-04
2^9	1.2e-15	2.0e-08	1e-7	3.1e-08	1.6e-05	2	1.7e-04	80	2.0e-03	80	1.1e-04
2^{10}	1.3e-15	3.3e-08	1e-7	5.0e-08	2.6e-05	2	3.3e-05	80	3.1e-03	80	2.6e-04
2^{11}	2.7e-10	2.5e-08	1e-7	2.6e-08	8.8e-06	2	7.2e-05	80	2.6e-03	80	2.6e-04
2^{12}	2.7e-10	3.5e-08	1e-7	3.9e-08	3.5e-06	2	8.7e-06	80	2.9e-03	80	3.9e-04
2^{13}	2.1e-10	3.6e-08	1e-7	3.7e-08	8.8e-06	2	2.5e-05	80	1.7e-03	80	5.0e-04
2^{14}	2.1e-10	3.3e-08	1e-7	4.4e-08	7.6e-06	2	7.1e-05	80	2.1e-03	80	6.6e-03
2^{15}	5.2e-10	3.1e-08	1e-7	3.6e-08	4.2e-05	2	1.4e-04	80	2.8e-03	80	2.6e-02
2^{16}	5.1e-10	3.3e-08	1e-7	2.9e-08	1.4e-05	2	1.2e-05	80	2.0e-03	80	1.7e-01

Table 4: Numerical comparison between different regularizers for kernel (5). $\lambda = 1E-8$. The iteration stops when $\|f^{k+1} - f^k\| < tol = 1E-10$.