1 March 16th, 2021

1.1 CG as a Direct Method

As proved before, GC will get the exact solution after t most n steps. In addition, the complexity per step is:

$$1 \text{ matrix-vector product } + \text{opeartions of } O(n)$$

Note that one matrix vector product is $O(m+n)$ where m is the number of nonzero entries in A. This means that the total computational cost is $O(mn+n^2)$ in the worse case.

- If A is the 1D Discrete Laplacian matrix, this is no better than Cholesky decomposition, which is $O(n)$.
- However if A is the 2D Discrete Laplacian, both are $O(n^2)$.

1.2 GC as an Iterative Method

CG can give a very accurate solution even if $k \ll n$.

Theorem 1.1

Assume A is SPD. Then $\{x_k\}$ generated by CG satisfies:

1. If A has only s distinct eigenvalues, then:
$$x_k = x_* \text{for all } k \geq s.$$
2. For a genera A: Let $\gamma = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}$ be the condition number, then we have:
$$\|x_k - x_*\|_A \leq 2\left(\frac{\sqrt{\gamma}-1}{\sqrt{\gamma}+1}\right)^k \|x_0 - x_*\|_A.$$
3. If eigenvalues of A satisfies:
$$0 < \lambda_1 \leq \ldots \leq \lambda_s \leq \alpha \leq \lambda_{s+1} \leq \ldots \leq \lambda_{n-t} \leq \beta \leq \lambda_{n-t+1} \leq \ldots \leq \lambda_n$$
Where α is close to β, (i.e. most eigenvalues are close together barring s small and t large outlying eigenvalues), then:
$$\|x_k - x_*\|_A \leq 2\left(\frac{\sqrt{\beta/\alpha}-1}{\sqrt{\beta/\alpha}+1}\right)^{k-s-t} \left(\max_{\lambda\in[\alpha,\beta]} \prod_{\ell\in\{1,\ldots,t\}\cup\{n-t+1,\ldots,n\}} \left|\frac{\lambda-\lambda_\ell}{\lambda_\ell}\right|\right)$$
Note that the right factor is a constant.

Corollary 1.2
From Theorem 1.1 (2), we have that the convergence speed depends on $O(\sqrt{\gamma})$, where as for steepest descent, it is $O(\gamma)$, meaning that the CG is much faster than steepest descent.

Example 1.3
If $A = (I + vv^T)$, then there are only two distinct eigenvalues, meaning that CG will converge in only two steps.

Proof. By the optimality of CG, we have:

$$\begin{aligned}
\|x_k - x_*\|_A &= \min_{x \in x_0 + K_k} \|x_* - x\|_A \\
&= \min_{c \in \mathbb{R}^k} \left\| x_* - (x_0 + \sum_{j=0}^{k-1} c_j A^j r_0) \right\|_A \\
&= \min_{c \in \mathbb{R}^k} \left\| (x_* - x_0) + \sum_{j=0}^{k-1} c_j A^{j+1} (x_* - x_0) \right\|_A \\
&= \min_{c \in \mathbb{R}^k} \left\| \left(I + \sum_{j=1}^{k} c_{j-1} A^j \right) (x_* - x_0) \right\|_A \\
&= \min_{p \in \mathbb{P}_k, p(0)=1} \|p(A)(x_* - x_0)\|_A \\
&\leq \left(\min_{p \in \mathbb{P}_k, p(0)=1} \|p(A)\|_A \right) \|(x_* - x_0)\|_A \\
&= \left(\min_{p \in \mathbb{P}_k, p(0)=1} \|p(A)\|_2 \right) \|(x_* - x_0)\|_A .
\end{aligned}$$

Where $\mathbb{P}_k$ is the set of polynomial of degree k.

Since A is symmetric, $p(A)$ is also symmetric. Thus, we have:

$$\begin{aligned}
\|x_k - x_*\|_A &\leq \left(\min_{p \in \mathbb{P}_k, p(0)=1} \|p(A)\|_2 \right) \|(x_* - x_0)\|_A \\
&= \left(\min_{p \in \mathbb{P}_k, p(0)=1} \max_{i \in \{1,\dots,n\}} |p(\lambda_i)| \right) \|(x_* - x_0)\|_A .
\end{aligned}$$

1. If A has only s distinct eigenvalues, say $\lambda_1, \dots, \lambda_s$, we have:

$$\min_{p \in \mathbb{P}_k, p(0)=1} \max_{i \in \{1,\dots,n\}} |p(\lambda_i)| \leq \max_{i \in \{1,\dots,n\}} |q(\lambda_i)| \quad \forall q \begin{cases} q \in \mathbb{P}_k \\ q(0) = 1 \end{cases}$$

Let us choose q by:

$$q(\lambda) = \prod_{i=1}^{s} \left(\frac{\lambda_i - \lambda}{\lambda_i} \right)$$

We have check that $q \in \mathbb{P}_s \subset \mathbb{P}_k$ and that $q(0) = 1$. With this, we have:

$$\begin{aligned} \min_{p\in\mathbb{P}_k, p(0)=1} \max_{i\in\{1,\dots,n\}} |p(\lambda_i)| &\leq \max_{i\in\{1,\dots,n\}} |q(\lambda_i)| \\ &= \max_{i\in\{i,\dots,s\}} |q(\lambda_i)| = 0. \end{aligned}$$

2. We relax the estimation by:

$$\begin{aligned} \|x_k - x_*\|_A &\leq \left(\min_{p\in\mathbb{P}_k, p(0)=1} \max_{i\in\{1,\dots,n\}} |p(\lambda_i)| \right) \|(x_* - x_0)\|_A \\ &\leq \left(\min_{p\in\mathbb{P}_k, p(0)=1} \max_{\lambda\in[\lambda_{\min},\lambda_{\max}]} |p(\lambda)| \right) \|(x_* - x_0)\|_A . \end{aligned}$$

Now we use a change of variable to estimate $\min\max|p(\lambda)|$. Define:

$$\mu = 2\frac{\lambda - \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}} - 1.$$

i.e. $\lambda = \lambda_{\min} \implies \mu = -1$, $\lambda = \lambda_{\max} \implies \mu = 1$. Thus, we estimate:

$$\min_{p\in\mathbb{P}_k, p(\mu_0)=1} \max_{\mu\in[-1,1]} |p(\mu)|$$

where $\mu_0 = 2\frac{-\lambda_{\min}}{\lambda_{\max}-\lambda_{\min}} - 1 = -\frac{\lambda_{\max}+\lambda_{\min}}{\lambda_{max}-\lambda_{\min}}$.

The solution of the minimax is given by the **Chebychev polynomial**.

Lemma 1.4

If $\mu_0 \neq [-1, 1]$, then:

$$\frac{C_k(\mu)}{C_k(\mu_0)} = \arg\min_{p\in\mathbb{P}_k, p(\mu_0)=1} \max_{\mu\in[-1,1]} |p(\mu)|$$

where:

$$C_k(\mu) = \begin{cases} \cos(k\cdot\arccos(\mu)) & \mu\in[-1,1] \\ \cosh(k\cdot\operatorname{arccosh}(\mu)) & \mu \geq 1 \\ (-1)^k\cosh(k\cdot\operatorname{arccosh}(-\mu)) & \mu \leq 1 \end{cases}$$

Proof. First we check that $C_k \in \mathbb{P}_k$. Indeed

$$\begin{aligned} C_0(\mu) &= 1 \in \mathbb{P}_0 \\ C_1(\mu) &= \mu \in \mathbb{P}_1. \end{aligned}$$

Also, by:

$$\begin{cases} \cos((k+1)\theta) + \cos((k-1)\theta) = 2\cos\theta\cos(k\theta) \\ \cosh((k+1)\theta) + \cosh((k-1)\theta) = 2\cosh\theta\cosh(k\theta) \end{cases}$$

Choosing $\theta = \arccos\mu$ if $|\mu| \leq 1$ or $\arccos h|\mu|$ if $|\mu| \geq 1$ and $k = k+1$, we have:

$$\begin{aligned} C_k(\mu) + C_{k-2}(\mu) &= 2\mu C_{k-1}(\mu) \\ \implies C_k(\mu) &= 2\mu C_{k-1}(\mu) - C_{k-2}(\mu) \in \mathbb{P}_k. \end{aligned}$$

This means that:

$$\frac{C_k(\mu)}{C_k(\mu_0)} \in \mathbb{P}_k \text{ and } \left.\frac{C_k(\mu)}{C_k(\mu_0)}\right|_{\mu=\mu_0} = 0.$$

Suppose there exists $q \neq \frac{C_k}{C_k(\mu_0)}$ s.t. $q \in \mathbb{P}_k,\ q(\mu_0) = 0$ and:

$$\max_{\mu\in[-1,1]} |q(\mu)| < \max_{\mu\in[-1,1]} \left|\frac{C_k(\mu)}{C_k(\mu_0)}\right| = \frac{1}{|C_k(\mu_0)|}$$

then consider:

$$f(\mu) = \frac{C_k(\mu)}{C_k(\mu_0)} - q(\mu) \in \mathbb{P}_k$$

Since:

$$C_k\left(\cos\frac{2j\pi}{k}\right) = \cos\left(k\cdot\arccos\left(\cos\frac{2j\pi}{k}\right)\right) = \cos\left(k\cdot\frac{2j\pi}{k}\right) = 1$$
$$C_k\left(\cos\frac{(2j+1)\pi}{k}\right) = \cos\left(k\cdot\arccos\left(\cos\frac{(2j+1)\pi}{k}\right)\right) = \cos\left(k\cdot\frac{(2j+1)\pi}{k}\right) = -1.$$

for any integer j s.t. $0 \le 2j, 2j+1 \le k$, and since $\cos 0, \cos\frac{\pi}{k}, \cos\frac{2\pi}{k}, \dots, \cos\frac{k\pi}{k}$ are $k+1$ distinct numbers in $[-1,1]$, WLOG, we assume $C_k(\mu_0) > 0$:

$$f(\mu)\begin{cases} \frac{1}{C_k(\mu_0)} - q(\mu) > 0 & \mu = \cos\frac{2j\pi}{k} \\ \frac{1}{C_k(\mu_0)} - q(\mu) < 0 & \mu = \cos\frac{(2j+1)\pi}{k} \end{cases}$$

Then f has at least k zeros, each in between $\left(\cos\frac{j\pi}{k}, \cos\frac{(j+1)\pi}{k}\right)$ for $j = 0, \dots, k$, and $f(\mu_0) = 0$ with $\mu_0 \notin [-1,1]$, meaning that $f \in \mathbb{P}_k$ has at least $k+1$ distinct zeros. However $f = 0$ is a contradiction. □

Continuing the convergence of CG, we have:

$$\begin{aligned}\|x_k - x_*\|_A &\le \max_{\mu\in[-1,1]}\left|\frac{C_k(\mu)}{C_k(\mu_0)}\right| \cdot \|x_0 - x_*\|_A \\ &= \frac{1}{|C_k(\mu_0)|}\|x_0 - x_*\|_A.\end{aligned}$$

It remains to give a lower bound of $|C_k(\mu_0)|$, with $\mu_0 < -1$.

Recall

$$\cosh(\theta) = \frac{e^\theta + e^{-\theta}}{2} \qquad \operatorname{arccosh}(x) = \ln(x + \sqrt{x^2-1})$$

and as such,

$$\begin{aligned}|C_k(\mu_0)| &= |\cosh(k\operatorname{arccosh}(-\mu_0))| \\ &= \frac{e^{k\ln(-\mu_0+\sqrt{\mu_0^2-1})} + e^{-k\ln(-\mu_0+\sqrt{\mu_0^2-1})}}{2} \\ &= \frac{1}{2}\left((\sqrt{\mu_0^2-1}-\mu_0)^k + (\sqrt{\mu_0^2-1}-\mu_0)^{-k}\right) \\ &\ge \frac{1}{2}(\sqrt{\mu_0-1}-\mu_0)^k.\end{aligned}$$

Note that:

$$\begin{aligned}\mu_0 &= -\frac{\lambda_{\max}+\lambda_{\min}}{\lambda_{\max}-\lambda_{\min}}\\ &= -\frac{\gamma+1}{\gamma-1}, \quad \gamma = \frac{\lambda_{\max}}{\lambda_{\max}}.\end{aligned}$$

Which gives us:

$$\begin{aligned}|C_k(\mu_0)| &\geq \frac{1}{2}(\sqrt{\mu_0-1}-\mu_0)^k\\ &= \frac{1}{2}\left(\sqrt{\left(\frac{\gamma+1}{\gamma-1}\right)^2+1}+\frac{\gamma+1}{\gamma-1}\right)^k\\ &= \frac{1}{2}\left(\sqrt{\frac{(\gamma+1)^2-(\gamma-1)^2}{(\gamma-1)^2}}+\frac{\gamma+1}{\gamma-1}\right)^k\\ &= \frac{1}{2}\left(\frac{2\sqrt{\gamma}+\gamma+1}{\gamma-1}\right)^k\\ &= \frac{1}{2}\left(\frac{(\sqrt{\gamma}+1)^2}{(\sqrt{\gamma}-1)(\sqrt{\gamma}+1)}\right)^k\\ &= \frac{1}{2}\left(\frac{\sqrt{\gamma}+1}{\sqrt{\gamma}-1}\right)^k.\end{aligned}$$

Thus:

$$\|x_0-x_*\|_A \leq 2\left(\frac{\sqrt{\gamma}+1}{\sqrt{\gamma}-1}\right)$$

For 3, we want to replace $\lambda_{\max}, \lambda_{\min}$ with α, β, meaning we construct a polynomial $q \in \mathbb{P}_k$ and $q(0)=1$ where:

$$q(\lambda) = \frac{C_{k-s-t}\left(2\frac{\lambda-\alpha}{\beta-\alpha}-1\right)}{C_{k-s-t}\left(-\frac{\beta+\alpha}{\beta-\alpha}\right)} \cdot \prod_{\ell\in\{1,\ldots,s\}\cup\{n-t+1,\ldots n\}}\left(\frac{\lambda_\ell-\lambda}{\lambda_\ell}\right)$$

Then:

$$\begin{aligned}\|x_k-x_*\|_A &\leq \min_{p\in}\max_{i=1}^{n}|p(\lambda_i)|\|x_0-x_*\|_A\\ &\leq \max_{i=1}^{n}|q(\lambda_i)|\|x_0-x_*\|_A.\end{aligned}$$

It remains to estimate $\max_{i=0}^{n}|q(\lambda_i)|$:

- When $i \in \{1, \dots, s\} \cup \{n-t+1, \dots, n\}$, $|q(\lambda_i)| = 1$.

$$\begin{aligned}
\max_{i=1}^{n} |q(\lambda_i)| &\le \max_{i \in \{s+1, \dots, n-t\}} |q(\lambda_i)| \\
&\le \max_{\lambda \in [\alpha, \beta]} |q(\lambda)| \\
&\le \max_{\lambda \in [\alpha, \beta]} \left| \frac{C_{k-s-t}\left(2\frac{\lambda-\alpha}{\beta-\alpha} - 1\right)}{C_{k-s-t}\left(-\frac{\beta+\alpha}{\beta-\alpha}\right)} \right| \cdot \max_{\lambda \in [\alpha, \beta]} \prod_{\ell \in \{1, \dots, s\} \cup \{n-t+1, \dots n\}} \left(\frac{\lambda_\ell - \lambda}{\lambda_\ell} \right) \\
&= 2 \left(\frac{\sqrt{\beta/\alpha} - 1}{\sqrt{\beta/\alpha} + 1} \right)^{k-s-t} \cdot \max_{\lambda \in [\alpha, \beta]} \prod_{\ell \in \{1, \dots, s\} \cup \{n-t+1, \dots n\}} \left(\frac{\lambda_\ell - \lambda}{\lambda_\ell} \right)
\end{aligned}$$

□

For a general SPD A in order to achieve an ϵ-solution, we want k to satisfy:

$$\begin{aligned}
\|x_k - x_*\|_A \le 2 \left(\frac{\sqrt{\gamma} - 1}{\sqrt{\gamma} + 1} \right)^k \|x_0 - x_*\|_A \le \epsilon \\
\implies k \ge \log \left(\frac{2\|x_0 - x_*\|_A}{\epsilon} \right) / \log \left(\frac{\sqrt{\gamma} + 1}{\sqrt{\gamma} - 1} \right).
\end{aligned}$$

Note that the numerator can be treated as a constant. Since:

$$\log \left(\frac{\sqrt{\gamma} + 1}{\sqrt{\gamma} - 1} \right) = \log \left(1 - \frac{2}{\sqrt{\gamma} - 1} \right) = O \left(\frac{1}{\sqrt{\gamma}} \right)$$

when γ is large. Thus, we have:

$$k \sim O(\log(1/\epsilon) \cdot \sqrt{\gamma}) = O(\sqrt{\gamma})$$

for a constant ϵ. Thus, if A is 2D discrete Laplacian, we have:

- Cholesky: $O(n^2)$
- Jacobi / G-S/Steepest Descent: $O(n^2)$
- CG for exact solution: $O(n^2)$
- CG for ϵ-solution: $O(n^{1.5})$

Thus, part 3 is used when the most of the eigenvalues of A are clustered with very few outliers. This will be useful in the preconditioning technique later.