The University of Texas at Austin
Department of Electrical and Computer Engineering

ECE 381V: Large-Scale Optimization II — Spring 2022

LECTURE 20

Caramanis & Mokhtari Wednesday, April 6, 2022

Goal: In this lecture, we talk about the Conjugate Gradient method and its convergence rate.

1 Linear systems with a PD matrix

Consider the problem of solving the following system of equations

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

where $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$ and $\mathbf{A} \in \mathbb{R}^{n\times n}$ is positive definite with bounded eigenvalues μ and L, where $0 < \mu \leq L$. This problem in particular is important when we are computing the direction of Newton's method, in which $\mathbf{A}$ corresponds to the objective function Hessian and $\mathbf{b}$ corresponds to the negative of the gradient direction.

It can be verified that the above problem is equivalent to the following quadratic problem

$$\min_{\mathbf{x}} \quad f(\mathbf{x}) := \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{b}^\top \mathbf{x}$$

In this lecture, we provide an efficient method that only requires gradient information and solves this problem efficiently.

Note that in this case, the objective function gradient equals the residual of the linear system:

$$\nabla f(\mathbf{x}_k) = \mathbf{A}\mathbf{x}_k - \mathbf{b} = \mathbf{r}_k$$

2 Conjugate Gradient Algorithm

The name Conjugate Gradient comes from the fact that in this method the directions selected for updating the iterates are conjugate with respect to matrix $\mathbf{A}$. In other words, if $\mathbf{p}_k$ refers to the descent direction selected at time k, i.e.,

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k$$

then we have

$$\mathbf{p}_i^\top \mathbf{A}\mathbf{p}_j = 0 \qquad \text{for } i \neq j$$

It further can be shown that these vectors $\mathbf{p}_1, \ldots, \mathbf{p}_k$ are linearly independent.

Why? Suppose

$$\mathbf{p}_j = \sum_{i\neq j} \alpha_i \mathbf{p}_i$$

Then by multiplying each side by $\mathbf{p}_l^\top \mathbf{A}$ for any $l \neq j$ we obtain that

$$0 = \mathbf{p}_l^\top \mathbf{A}\mathbf{p}_j = \mathbf{p}_l^\top \mathbf{A} \sum_{i\neq j} \alpha_i \mathbf{p}_i = \alpha_l \mathbf{p}_l^\top \mathbf{A}\mathbf{p}_l = \alpha_l \|\mathbf{p}\|_{\mathbf{A}}$$

and hence, $\alpha_l = 0$ for any $l \neq j$.

2.1 Why conjugacy is important?

It is important because we can solve the problem exactly after at most n iterations!

To prove the above claim suppose $\mathbf{p}_1, \ldots, \mathbf{p}_k$ are all conjugate wrt $\mathbf{A}$ and they are linearly independent. Further update the iterates according to the update

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k$$

where

$$\alpha_k = -\frac{\nabla f(\mathbf{x}_k)^\top \mathbf{p}_k}{\mathbf{p}_k^\top \mathbf{A} \mathbf{p}_k} = -\frac{\mathbf{r}_k^\top \mathbf{p}_k}{\mathbf{p}_k^\top \mathbf{A} \mathbf{p}_k}$$

Now note that after n iterations $\mathbf{p}_0, \ldots, \mathbf{p}_{n-1}$ create a basis for $\mathbb{R}^n$, hence there exits $\sigma_0, \ldots, \sigma_{n-1}$ such that

$$\mathbf{x}^* - \mathbf{x}_0 = \sum_{i=0}^{n-1} \sigma_i \mathbf{p}_i$$

Now by multiplying both sides by $\mathbf{p}_j^\top \mathbf{A}$ we obtain that

$$\sigma_j = \frac{\mathbf{p}_j^\top \mathbf{A}(\mathbf{x}^* - \mathbf{x}_0)}{\mathbf{p}_j^\top \mathbf{A} \mathbf{p}_j}$$

Hence, using the fact that $\mathbf{A}(\mathbf{x}^* - \mathbf{x}_0) = \mathbf{b} - \mathbf{A}\mathbf{x}_0 = -\mathbf{r}_0$ we have

$$\mathbf{x}^* \ = \ \mathbf{x}_0 + \sum_{j=0}^{n-1} \frac{\mathbf{p}_j^\top \mathbf{A}(\mathbf{x}^* - \mathbf{x}_0)}{\mathbf{p}_j^\top \mathbf{A} \mathbf{p}_j} \mathbf{p}_j \ = \ \mathbf{x}_0 - \sum_{j=0}^{n-1} \frac{\mathbf{p}_j^\top \mathbf{r}_0}{\mathbf{p}_j^\top \mathbf{A} \mathbf{p}_j} \mathbf{p}_j$$

On the other hand we have

$$\mathbf{x}_n = \mathbf{x}_0 - \sum_{j=0}^{n-1} \frac{\mathbf{r}_j^\top \mathbf{p}_j}{\mathbf{p}_j^\top \mathbf{A} \mathbf{p}_j} \mathbf{p}_j$$

Hence, we have $\mathbf{x}^* = \mathbf{x}_n$ if we show that for any $0 \le j \le n-1$ we have

$$\mathbf{r}_j^\top \mathbf{p}_j = \mathbf{r}_0^\top \mathbf{p}_j$$

This is easy to prove as

$$\mathbf{r}_j^\top \mathbf{p}_j = \mathbf{r}_0^\top \mathbf{p}_j \iff (\mathbf{r}_j - \mathbf{r}_0)^\top \mathbf{p}_j = 0 \iff (\mathbf{A}\mathbf{x}_j - \mathbf{A}\mathbf{x}_0)^\top \mathbf{p}_j = 0 \iff (\mathbf{x}_j - \mathbf{x}_0)^\top \mathbf{A} \mathbf{p}_j = 0$$

where the last condition holds as $\mathbf{x}_j - \mathbf{x}_0 = \sum_{i=0}^{j-1} \mathbf{p}_i$ and each term is conjugate to $\mathbf{p}_j$ wrt $\mathbf{A}$. Hence, we can conclude that $\mathbf{x}_n = \mathbf{x}^*$.

2.2 How should we select vectors $\mathbf{p}_k$?

The first direction is $\mathbf{p}_0 = -\mathbf{r}_0$, which is the gradient direction. Then for $k \ge 1$ we have

$$\mathbf{p}_k = -\mathbf{r}_k \ + \ \frac{\mathbf{r}_k^\top \mathbf{A} \mathbf{p}_{k-1}}{\mathbf{p}_{k-1}^\top \mathbf{A} \mathbf{p}_{k-1}} \ \mathbf{p}_{k-1}$$

This update is selected such that $\mathbf{p}_{k-1}^\top \mathbf{A} \mathbf{p}_k = 0$ holds. (We still don't know if $\mathbf{p}_1, \ldots, \mathbf{p}_k$ are all conjugate with respect to $\mathbf{A}$! But, we will prove it!)

2.3 Conjugate Gradient: The general version

Set $\mathbf{x}_0 = \mathbf{0}$ and $\mathbf{p}_0 = -\mathbf{r}_0 = \mathbf{b}$. Further, set $k = 0$

- Step 1: Compute stepsize $\alpha_k = -\frac{\mathbf{r}_k^\top \mathbf{p}_k}{\mathbf{p}_k^\top \mathbf{A}\mathbf{p}_k}$
- Step 2: Update the iterate: $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k$
- Step 3: Update the direction $\mathbf{p}_k = -\mathbf{r}_k \; + \; \frac{\mathbf{r}_k^\top \mathbf{A}\mathbf{p}_{k-1}}{\mathbf{p}_{k-1}^\top \mathbf{A}\mathbf{p}_{k-1}} \; \mathbf{p}_{k-1}$
- Step 4: Set $k := k + 1$ and go to Step 1.

3 Krylov subspaces

In this section, we discuss Krylov subspaces and Krylov sequences which are closely related to the Conjugate Gradient algorithm.
Consider a sequence of subspaces defined as

$$\mathcal{K}_0 = \{\mathbf{0}\}, \qquad\qquad \mathcal{K}_k = \mathit{span}\{\mathbf{b}, \mathbf{A}\mathbf{b}, \ldots, \mathbf{A}^{k-1}\mathbf{b}\} \text{ for } k \geq 1.$$

It can be easily verified that this sequence satisfies the following conditions:

- $\mathcal{K}_0 \subseteq \mathcal{K}_1 \subseteq \mathcal{K}_2, \ldots,$
- $\dim(\mathcal{K}_{k+1}) - \dim(\mathcal{K}_k) \in \{0, 1\}$
- If $\mathcal{K}_{k+1} = \mathcal{K}_k$ then $\mathcal{K}_i = \mathcal{K}_k$ for all $i \geq k$

Remark 1. *$\mathcal{K}_n$ may not be $\mathbb{R}^n$. For instance, for $\mathbf{A} = \mathbf{I}$ we have $\{\mathbf{b}\} = \mathcal{K}_1 = \mathcal{K}_2 = \cdots = \mathcal{K}_n = \ldots$*

Theorem 1. *Consider the defintion $\mathcal{K}_n = \{\mathbf{b}, \mathbf{A}\mathbf{b}, \ldots, \mathbf{A}^{n-1}\mathbf{b}\}$. Then, we have $\mathbf{A}^{-1}\mathbf{b} \in \mathcal{K}_n$*

Proof. Note that based on Cayley–Hamilton theorem we have

$$P_{\mathbf{A}}(\mathbf{A}) = \mathbf{A}^n + a_1 \mathbf{A}^{n-1} + \cdots + a_{n-1}\mathbf{A} + a_n \mathbf{I} = \mathbf{0}$$

where

$$P_{\mathbf{A}}(\lambda) = \det(\lambda \mathbf{I} - \mathbf{A}) = \lambda^n + a_1 \lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$$

This observation shows that by multiplying both sides from right by $\mathbf{A}^{-1}\mathbf{b}$ we have

$$\mathbf{A}^{-1}\mathbf{b} = \left(\frac{-1}{a_n}\right) (\mathbf{A}^{n-1}\mathbf{b} + a_1 \mathbf{A}^{n-2}\mathbf{b} + \cdots + a_{n-1}\mathbf{b})$$

□

3.1 Krylov sequences

Now consider the sequence of iterates generated according to the update

$$\mathbf{x}_k = \underset{\mathbf{x} \in \mathcal{K}_k}{\operatorname{argmin}} f(\mathbf{x})$$

From what we discussed it can be easily verified that

$$\mathbf{x}_n = \mathbf{A}^{-1}\mathbf{b}$$

4 Properties of the Conjugate Gradient Algorithm

Based on the above discussion we can think of the update of CG as

$$\mathbf{x}_k = \operatorname*{argmin}_{\mathbf{x} \in \mathcal{K}_k} f(\mathbf{x})$$

where here $\mathcal{K}_k = span\{\mathbf{b}, \dots, \mathbf{A}^{k-1}\mathbf{b}\} = span\{\mathbf{r}_0, \dots, \mathbf{A}^{k-1}\mathbf{r}_0\}$. Before proving this result, we establish the following results for the iterates of CG.

Theorem 2. *Consider the iterates of CG. Then the following results hold:*

- *(a)* $\mathbf{r}_{k+1} = \mathbf{r}_k + \alpha_k \mathbf{A}\mathbf{p}_k$
- *(b)* $span\{\mathbf{r}_0, \dots, \mathbf{r}_k\} = span\{\mathbf{r}_0, \dots, \mathbf{A}^k\mathbf{r}_0\}$
- *(c)* $span\{\mathbf{p}_0, \dots, \mathbf{p}_k\} = span\{\mathbf{r}_0, \dots, \mathbf{A}^k\mathbf{r}_0\}$
- *(d)* $\mathbf{r}_k^\top \mathbf{r}_j = 0$ *for* $j < k$
- *(e)* $\mathbf{p}_k^\top \mathbf{A}\mathbf{p}_j = 0$ *for* $k \neq j$

Proof. Here are the proofs:

Proof of (a). Note that we can easily prove (a) using the fact that

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k \quad \Rightarrow \quad \mathbf{A}\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \alpha_k \mathbf{A}\mathbf{p}_k \quad \Rightarrow \quad \mathbf{A}\mathbf{x}_{k+1} - \mathbf{b} = \mathbf{A}\mathbf{x}_k - \mathbf{b} + \alpha_k \mathbf{A}\mathbf{p}_k$$

Proof of (b) and (c). Next we prove (b) and (c) together using induction.
Base Step. For $k = 0$ both (b) and (c) are trivially satisfied.
Induction Step. Suppose $span\{\mathbf{r}_0, \dots, \mathbf{r}_k\} = span\{\mathbf{r}_0, \dots, \mathbf{A}^k\mathbf{r}_0\}$ and $span\{\mathbf{p}_0, \dots, \mathbf{p}_k\} = span\{\mathbf{r}_0, \dots, \mathbf{A}^k\mathbf{r}_0\}$. Then, we have

$$\mathbf{A}\mathbf{p}_k \in span\{\mathbf{A}\mathbf{r}_0, \dots, \mathbf{A}^{k+1}\mathbf{r}_0\} \;\Rightarrow\; \mathbf{r}_{k+1} \in span\{\mathbf{A}\mathbf{r}_0, \dots, \mathbf{A}^{k+1}\mathbf{r}_0\} \;\Rightarrow\; span\{\mathbf{r}_0, \dots, \mathbf{r}_{k+1}\} \subset span\{\mathbf{r}_0, \dots, \mathbf{A}^{k+1}\mathbf{r}_0$$

Further, we have

$$\mathbf{A}^{k+1}\mathbf{r}_0 \in span\{\mathbf{A}\mathbf{p}_0, \dots, \mathbf{A}\mathbf{p}_k\} \;\Rightarrow\; \mathbf{A}^{k+1}\mathbf{r}_0 \in span\{\mathbf{r}_0, \dots, \mathbf{r}_{k+1}\}$$

where we used the fact that $span\{\mathbf{A}\mathbf{p}_0, \dots, \mathbf{A}\mathbf{p}_k\} \subset span\{\mathbf{r}_0, \dots, \mathbf{r}_{k+1}\}$. By combining these results we obtain that

$$span\{\mathbf{r}_0, \dots, \mathbf{r}_{k+1}\} = span\{\mathbf{r}_0, \dots, \mathbf{A}^{k+1}\mathbf{r}_0\}$$

Now to complete the induction step for (c) we use the fact that

$$span\{\mathbf{p}_0, \dots, \mathbf{p}_{k+1}\} = span\{\mathbf{p}_0, \dots, \mathbf{p}_k, \mathbf{r}_{k+1}\} = span\{\mathbf{r}_0, \dots, \mathbf{A}^k\mathbf{r}_0, \mathbf{r}_{k+1}\} = span\{\mathbf{r}_0, \dots, \mathbf{r}_k, \mathbf{r}_{k+1}\}$$

Now finally by using (b) for $k + 1$ we can conclude that

$$span\{\mathbf{p}_0, \dots, \mathbf{p}_{k+1}\} = span\{\mathbf{r}_0, \dots, \mathbf{A}^{k+1}\mathbf{r}_0\}$$

and the induction proof for part (c) is done.

Proof of (d) and (e). Exercise. $\square$

5 Optimality of Conjugate Gradient

Note that we know that $\mathbf{x}_k \in \mathbf{x}_0 + \mathcal{K}_k$. Now we need to show it is the minimizer of the quadratic function over this subspace.

Theorem 3. *The sequence of iterates generated by CG are equivalent to the sequence*

$$\mathbf{x}_k = \underset{\mathbf{x}-\mathbf{x}_0 \in \mathcal{K}_k}{\operatorname{argmin}} f(\mathbf{x})$$

where $\mathcal{K}_k = \mathit{span}\{\mathbf{b}, \ldots, \mathbf{A}^{k-1}\mathbf{b}\} = \mathit{span}\{\mathbf{r}_0, \ldots, \mathbf{A}^{k-1}\mathbf{r}_0\}$.

Proof. To prove this claim, first define $\mathbf{e}_k = \mathbf{x}_k - \mathbf{x}^*$ and its weighted norm as

$$\|\mathbf{e}_k\|_{\mathbf{A}}^2 := \mathbf{e}_k^\top \mathbf{A} \mathbf{e}_k = (\mathbf{x}_k - \mathbf{x}^*)^\top \mathbf{A} (\mathbf{x}_k - \mathbf{x}^*) = 2f(\mathbf{x}_k)$$

Now we show that for any other $\hat{\mathbf{x}}$ such that $\hat{\mathbf{x}} - \mathbf{x}_0 \in \mathcal{K}_k$ we have $\|\mathbf{e}_k\|_{\mathbf{A}}^2 \leq \|\hat{\mathbf{e}}\|_{\mathbf{A}}^2$. To do so, note that $\hat{\mathbf{x}} = \mathbf{x}_k - \Delta\mathbf{x}$ which implies that $\hat{\mathbf{e}} = \mathbf{e}_k + \Delta\mathbf{x}$. Hence, we have

$$\|\hat{\mathbf{e}}\|_{\mathbf{A}}^2 = \|\mathbf{e}_k\|_{\mathbf{A}}^2 + 2\mathbf{e}_k^\top \mathbf{A} \Delta\mathbf{x} + \Delta\mathbf{x}^\top \mathbf{A} \Delta\mathbf{x} = \|\mathbf{e}_k\|_{\mathbf{A}}^2 + \Delta\mathbf{x}^\top \mathbf{A} \Delta\mathbf{x}$$

where the last equality follows from the fact that

$$\mathbf{e}_k^\top \mathbf{A} \Delta\mathbf{x} = \mathbf{r}_k^\top \Delta\mathbf{x} = 0$$

since $\Delta\mathbf{x} \in \mathcal{K}_k$ and we know $\mathbf{r}_k$ is orthogonal to any point in this subspace. Finally, the proof is complete using the fact that $\Delta\mathbf{x}^\top \mathbf{A} \Delta\mathbf{x} > 0$ for any $\Delta\mathbf{x} \neq \mathbf{0}$.

□

It is also obvious that

$$\|\mathbf{e}_k\|_{\mathbf{A}}^2 \leq \|\mathbf{e}_{k-1}\|_{\mathbf{A}}^2$$

since $\mathcal{K}_{k-1} \subseteq \mathcal{K}_k$.

6 Convergence Rate

Note that $\mathbf{x}_{k+1} = \operatorname{argmin}_{\mathbf{x} \in \mathcal{K}_{k+1}} f(\mathbf{x})$. In another words, $\mathbf{x}_{k+1}$ is the minimizer of the quadratic loss for the set of points that can be written as

$$\hat{\mathbf{x}} = \mathbf{x}_0 + \mathit{span}\{\mathbf{r}_0, \ldots, \mathbf{A}^k \mathbf{r}_0\}$$

Therefore, if we consider any polynomial of the form

$$P_k(\mathbf{A})(\gamma_1, \ldots, \gamma_n) = \gamma_0 \mathbf{I} + \gamma_1 \mathbf{A} + \cdots + \gamma_k \mathbf{A}^k$$

Then, any point in $\mathit{span}\{\mathbf{r}_0, \ldots, \mathbf{A}^k \mathbf{r}_0\}$ can be written as $P_k(\mathbf{A})(\gamma_1, \ldots, \gamma_n)\mathbf{r}_0$ for some $(\gamma_1, \ldots, \gamma_n)$. Hence, we can say that

$$\mathbf{x}_{k+1} = P_k(\mathbf{A})(\gamma_1^*, \ldots, \gamma_n^*)\mathbf{r}_0 := P_k(\mathbf{A})^* \mathbf{r}_0$$

where

$$\min_{(\gamma_1, \ldots, \gamma_n)} \|\mathbf{x}_0 + P_k(\mathbf{A})(\gamma_1, \ldots, \gamma_n)\mathbf{r}_0 - \mathbf{x}^*\|_{\mathbf{A}}^2$$

Further we have

$$\mathbf{x}_{k+1} - \mathbf{x}^* = \mathbf{x}_0 + P_k(\mathbf{A})^*\mathbf{r}_0 - \mathbf{x}^* = \mathbf{x}_0 + P_k(\mathbf{A})^*\mathbf{A}(\mathbf{x}_0 - \mathbf{x}^*) - \mathbf{x}^* = (\mathbf{I} + P_k(\mathbf{A})^*\mathbf{A})(\mathbf{x}_0 - \mathbf{x}^*)$$

Now using the decompositions

$$\mathbf{A} = \sum_{i=1}^{n} \lambda_i \mathbf{v}_i \mathbf{v}_i^\top, \qquad \mathbf{x}_0 - \mathbf{x}^* = \sum_{i=1}^{n} \zeta_i \mathbf{v}_i$$

and using the fact that $P_k(\mathbf{A})\mathbf{v}_i = P_k(\lambda_i)\mathbf{v}_i$ we have

$$\mathbf{x}_{k+1} - \mathbf{x}^* = \sum_{i=1}^{n} (1 + \lambda_i P_k(\lambda_i)^*)\zeta_i \mathbf{v}_i$$

using the fact that $\|\mathbf{z}\|_{\mathbf{A}}^2 = \mathbf{z}^\top \mathbf{A} \mathbf{z} = \sum_{i=1}^{n} \lambda_i (\mathbf{v}_i \mathbf{z})^2$ we have

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\|_{\mathbf{A}}^2 = \sum_{i=1}^{n} \lambda_i (1 + \lambda_i P_k(\lambda_i)^*)^2 \zeta_i^2$$

Since the polynomial for CG is optimal we have

$$\begin{aligned}
\|\mathbf{x}_{k+1} - \mathbf{x}^*\|_{\mathbf{A}}^2 &= \min_{P_k} \sum_{i=1}^{n} \lambda_i (1 + \lambda_i P_k(\lambda_i))^2 \zeta_i^2 \\
&\leq \min_{P_k} \max_{i=1,\ldots,n} (1 + \lambda_i P_k(\lambda_i))^2 \sum_{i=1}^{n} \lambda_i \zeta_i^2 \\
&\leq \min_{P_k} \max_{i=1,\ldots,n} (1 + \lambda_i P_k(\lambda_i))^2 \|\mathbf{x}_0 - \mathbf{x}^*\|_{\mathbf{A}}^2
\end{aligned}$$

Hence the contraction factor after $k+1$ iterations is denoted by

$$\min_{P_k} \max_{i=1,\ldots,n} (1 + \lambda_i P_k(\lambda_i))^2$$

Theorem 4. *If $\mathbf{A}$ has r distinct eigenvalues then CG converges after at most r iterations.*

Proof. Suppose those eigenvalues are denoted by $\lambda_1, \ldots, \lambda_r$. One can easily verify that the polynomial

$$P_k(\lambda) = \prod_{i=1,\ldots,r} \left(1 - \frac{\lambda}{\lambda_i}\right)$$

is a valid polynomial for $k \geq r$ and it is zero at all eigenvalues. □

Theorem 5. *After k iterations for CG method we have*

$$\|\mathbf{x}_k - \mathbf{x}^*\|_{\mathbf{A}}^2 \leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}\right)^k \|\mathbf{x}_0 - \mathbf{x}^*\|_{\mathbf{A}}^2$$

Proof. Exercise. □