MATH 6222 LECTURE NOTE 4:
CONJUGATE GRADIENT METHOD

JINGRONG WEI

ABSTRACT. This note gives introduction to the conjudate gradient method. Main references are [3, Chapter 5] and [1].

CONTENTS

The conjugate gradient method developed by Hestenes and Stiefel in 1950s [2] is an iterative method for solving a linear system of equations

$$Ax = b, \tag{1}$$

where A is a symmetric positive definite (SPD) operator defined on an n-dimensional Hilbert space $\mathcal{V}$ with inner product $(\cdot,\cdot)$, and $b \in \mathcal{V}$. The problem (1) can be stated equivalently as the following minimization problem:

$$\min_x f(x) = \frac{1}{2}(x, Ax) - (b, x), \tag{2}$$

As f is strongly convex, the global minimizer x^* exists and is unique and satisfies $\nabla f(x^*) = 0$, which is exactly equation (1). This equivalence will allow us to interpret the conjugate gradient method either as an algorithm for solving linear systems or as a technique for minimizing convex quadratic functions. For future reference, we note that the gradient of f equals the residual of the linear system, that is,

$$\nabla f(x) = Ax - b := r(x).$$

We use $(\cdot,\cdot)_A$ for the inner product introduced by the SPD operator A :

$$(x, y)_A = (Ax, y) = (x, Ay) = x^\top Ay = y^\top Ax,$$

which induces a norm $\|x\|_A = \sqrt{(x, x)_A}$.

Date: February 5, 2025.

1. Conjugate Direction Methods

A set of nonzero vectors $\{p_i\}$ is said to be *conjugate* with respect to the symmetric positive definite matrix A if

$$(p_i, p_j)_A = p_i^\top A p_j = 0, \quad \forall\, i \neq j, \tag{3}$$

which is the same to say p_i and p_j are A-orthogonal for $i \neq j$.

We shall derive conjugate direction method from the A-orthogonal projection to subspaces. Given a vector $x \in \mathcal{V}$, the A-orthogonal projection of x to a subspace $S \subseteq \mathcal{V}$ is a vector in S, denoted by $\operatorname{Proj}_S^A x$, by the relation

$$\left(\operatorname{Proj}_S^A x, y\right)_A = (x, y)_A, \quad \forall y \in S.$$

Suppose we can find an A-orthogonal basis, i.e.

$$\mathcal{V}_k = \operatorname{span}\left\{p_0, p_1, \cdots, p_k\right\},$$

the projection can be found component by component

$$\operatorname{Proj}_{\mathcal{V}_k}^A\left(x^* - x_0\right) = \sum_{i=0}^{k} \alpha_i p_i, \quad \alpha_i = \frac{\left(x^* - x_0, p_i\right)_A}{\left(p_i, p_i\right)_A}, \quad \text{for } i = 0, \cdots k.$$

Then the approximation

$$x_{k+1} = x_0 + \sum_{i=0}^{k} \alpha_i p_i$$

is the 'best' approximation in the sense that

$$\operatorname{Proj}_{\mathcal{V}_k}^A(x^* - x_{k+1}) = \operatorname{Proj}_{\mathcal{V}_k}^A(x^* - x_0) - \sum_{i=0}^{k} \alpha_i p_i = 0, \tag{4}$$

that is, x_{k+1} is the A-orthogonal projection of x^* onto $\mathcal{V}_k$. To summarize, we give the conjugate direction method in Algorithm 1.

Algorithm 1 Conjugate direction method for solving $Ax = b$.

1: **Parameters:** $x_0 \in \mathbb{R}^n$ **and a set of conjugate directions** $\{p_0, p_1, \dots, p_{n-1}\}$.
2: **for** $k = 0, 1, \dots, n-1$ **do**
3: $\quad r_k = Ax_k - b$
4: $\quad x_{k+1} = x_k + \alpha_k p_k$ with $\alpha_k = -\frac{r_0^\top p_k}{p_k^\top A p_k} = -\frac{r_k^\top p_k}{p_k^\top A p_k}$
5: **end for**
6: **return** x_n

Theorem 1.1. *For any $x_0 \in \mathbb{R}^n$, the sequence $\{x_k\}$ generated by the conjugate direction algorithm 1 converges to the solution x^* of the linear system (1) in at most n steps.*

Proof. It is not hard to check that $\{p_i\}_{i=0}^n$ are linearly independent and they span the whole space $\mathbb{R}^n$. By the orthogonality, we can write:

$$x^* - x_0 = \alpha_0 p_0 + \alpha_1 p_1 + \cdots + \alpha_{n-1} p_{n-1}.$$

□

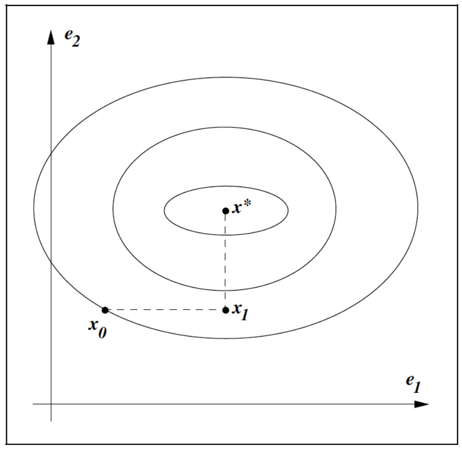

FIGURE 1. Successive minimizations along the coordinate directions find the minimizer of a quadratic with a diagonal Hessian in n iterations.

The coefficient α_k can be simplified using

$$x_k = x_0 + \alpha_0 p_0 + \alpha_1 p_1 + \cdots + \alpha_{k-1} p_{k-1}.$$

By premultiplying this expression by $p_k^\top A$ and using the conjugacy property, we have that

$$p_k^\top A\,(x_k - x_0) = 0$$

and therefore

$$p_k^\top A\,(x^* - x_0) = p_k^\top A\,(x^* - x_k) = p_k^\top\,(b - Ax_k) = -p_k^\top r_k.$$

There is a simple interpretation of the properties of conjugate directions. If the matrix A is diagonal, the contours of the function $f(\cdot)$ are ellipses whose axes are aligned with the coordinate directions, as illustrated in Figure 1. We can find the minimizer of this function by performing one-dimensional minimization along the coordinate directions.

When A is not diagonal, its contours are still elliptical, but they are usually no longer aligned with the coordinate directions. We can, however, recover the nice behavior of Figure 1 if we transform the problem to make A diagonal and then minimize along the coordinate directions. Suppose we transform the problem by defining new variables $\hat{x}$ as

$$\hat{x} = S^{-1}x,$$

where S is the $n \times n$ matrix defined by

$$S = \begin{bmatrix} p_0 & p_1 & \cdots & p_{n-1} \end{bmatrix}.$$

The quadratic f defined by (2) now becomes

$$\hat{f}(\hat{x}) := f(S\hat{x}) = \frac{1}{2}\hat{x}^\top \left(S^\top A S\right)\hat{x} - \left(S^\top b\right)^\top \hat{x}.$$

By the conjugacy property (3), the matrix $S^\top A S$ is diagonal, so we can find the minimizing value of $\hat{f}$ by performing n one-dimensional minimizations along the coordinate directions of $\hat{x}$. Because of the relation $x = \hat{S}\hat{x}$, however, the i th coordinate direction in $\hat{x}$-space corresponds to the direction p_i in x-space. Hence, the coordinate search strategy applied to $\hat{f}$ is equivalent to the conjugate direction algorithm on f.

Theorem 1.2 (Expanding Subspace Minimization)**.** *Let $x_0 \in \mathbb{R}^n$ be any starting point and suppose that the sequence $\{x_k\}$ is generated by the conjugate direction algorithm 1. Then*

$$r_k^\top p_i = 0, \quad i = 0, 1, \ldots, k-1 \tag{5}$$

and x_k is the minimizer of $f(x) = \frac{1}{2}x^\top Ax - b^\top x$ over the set

$$\{x \mid x - x_0 \in \mathcal{V}_{k-1} = \operatorname{span}\{p_0, p_1, \ldots, p_{k-1}\}\} \tag{6}$$

Proof. We shall show r_k is orthogonal to $\mathcal{V}_{k-1}$. Notice that $x^* - x_k$ is A-orthogonal to $\mathcal{V}_{k-1}$ by (4), for all $v \in \mathcal{V}_{k-1}$,

$$(x^* - x_k, v)_A = (b - Ax_k, v) = -(r_k, v) = 0,$$

which implies (5).

To complete the proof, we show that a point $\tilde{x}$ minimizes f over the set (6) if and only if $r(\tilde{x})^\top p_i = 0$, for each $i = 0, 1, \ldots, k-1$. Let us define $h(\sigma) = \phi\left(x_0 + \sigma_0 p_0 + \cdots + \sigma_{k-1}p_{k-1}\right)$, where $\sigma = (\sigma_0, \sigma_1, \ldots, \sigma_{k-1})^\top$. Since $h(\sigma)$ is a strictly convex quadratic, it has a unique minimizer σ^* that satisfies

$$\frac{\partial h\left(\sigma^*\right)}{\partial \sigma_i} = 0, \quad i = 0, 1, \ldots, k-1,$$

and

$$\tilde{x} = x_0 + \sigma_0^* p_0 + \sigma_1^* p_2 + \cdots + \sigma_{k-1}^* p_{k-1}.$$

By the chain rule, this equation implies that

$$\nabla\phi\left(\tilde{x}\right)^\top p_i = r(\tilde{x})^\top p_i = 0, \quad i = 0, 1, \ldots, k-1.$$

□

2. Conjugate Gradient Method

The conjugate gradient (CG) method is a conjugate direction method with a very special property: In generating its set of conjugate vectors, it can compute a new vector p_k by using only the previous vector p_{k-1}. It does not need to know all the previous elements $p_0, p_1, \ldots, p_{k-2}$ of the conjugate set; p_k is automatically conjugate to these vectors. This remarkable property implies that the method requires little storage and computation.

In the conjugate gradient method, each direction p_k is the conjugate direction from the negative residual $-r_k$ (which is the steepest descent direction for the function f) and the previous direction p_i's. If $r_k = 0$, which means $x_k = x^*$ is the solution, then we stop. Otherwise we expand the subspace to a larger one $\mathcal{V}_{k+1} = \operatorname{span}\{p_0, p_1, \cdots, p_k, -r_{k+1}\}$.

Then apply Gram-Schmit process to make new added vector $-r_{k+1}$ to be A-orthogonal to others. The new conjugate direction is

$$p_{k+1} = -r_{k+1} + \sum_{i=0}^{k} \beta_i p_i, \quad \beta_i = \frac{(r_{k+1}, p_i)_A}{(p_i, p_i)_A}$$

The magic of CG algorithm is that only β_k is needed due to the orthogonality we shall explore now.

Lemma 2.1. $r_k^\top p_k = -r_k^\top r_k, k = 0, 1, \ldots.$

Lemma 2.2. *The residual r_{k+1} is A-orthogonal to $\mathcal{V}_{k-1}$.*

Proof. If the algorithm stops at the k-th step, then $r_{k+1} = 0$ and the statement is true. Otherwise the algorithm does not stop at the k-th step implies $r_i \neq 0$ and consequently $\alpha_i \neq 0$ for $i \leq k-1$. By the recursive formula for the residual $r_{i+1} = r_i + \alpha_i A p_i$. As $\alpha_i \neq 0$, we get $Ap_i \in \operatorname{span}\{r_i, r_{i+1}\} \subset \mathcal{V}_k$ for $0 \leq i \leq k-1$. Since we have proved $r_{k+1} \perp \mathcal{V}_k$ (Theorem 1.2), we get $(r_{k+1}, p_i)_A = (r_{k+1}, Ap_i) = 0$ for $0 \leq i \leq k-1$, i.e. r_{k+1} is A-orthogonal to $\mathcal{V}_{k-1}$. □

Therefore, we conclude that

$$p_{k+1} = -r_{k+1} + \beta_k p_k, \quad \beta_k = \frac{r_{k+1}^\top A p_k}{p_k^\top A p_k}.$$

The conjugate gradient method is summarized in Algorithm 2.

Algorithm 2 Conjugate gradient method for solving $Ax = b$.

Parameters: $x_0 \in \mathbb{R}^n$
Set $r_0 = Ax_0 - b, p_0 = -r_0$
for $k = 0, 1, \dots$ **do**
  $x_{k+1} = x_k + \alpha_k p_k$ with $\alpha_k = -\dfrac{r_k^\top p_k}{p_k^\top A p_k} = \dfrac{r_k^\top r_k}{p_k^\top A p_k}$
  $r_{k+1} = Ax_{k+1} - b$
  $p_{k+1} = -r_{k+1} + \beta_k p_k$ with $\beta_k = \dfrac{r_{k+1}^\top A p_k}{p_k^\top A p_k} = \dfrac{r_{k+1}^\top r_{k+1}}{r_k^\top r_k}$
end for
return x_{k+1}

Theorem 2.3 (Properties of Conjugate Gradient Method)**.** *Suppose that the kth iterate generated by the conjugate gradient method is not the solution point x^*. The following properties hold:*

(7a)
$$\mathcal{V}_k = \operatorname{span}\{p_0, p_1, \dots, p_k\} = \operatorname{span}\{r_0, r_1, \dots, r_k\} = \operatorname{span}\left\{r_0, Ar_0, \dots, A^k r_0\right\},$$
(7b)
$$r_k^\top r_i = 0, \quad i = 0, 1, \dots, k-1.$$

Therefore, the sequence $\{x_k\}$ converges to x^* in at most n steps.

Proof. $\mathcal{V}_k = \operatorname{span}\{r_0, r_1, \dots, r_k\}$ is straightforward by construction.

We prove by induction. (7a) holds trivially for $k = 0$. Assuming now that (7a) is true for some k (the induction hypothesis), we show that they continue to hold for $k+1$.

Because of the induction hypothesis,

$$\{r_k, p_k\} \in \operatorname{span}\left\{r_0, Ar_0, \dots, A^k r_0\right\},$$

we obtain

$$r_{k+1} = r_k + \alpha_k A p_k \in \operatorname{span}\left\{r_0, Ar_0, \dots, A^{k+1} r_0\right\}.$$

Therefore, we conclude that

$$\operatorname{span}\{r_0, r_1, \dots, r_k, r_{k+1}\} \subset \operatorname{span}\left\{r_0, Ar_0, \dots, A^{k+1} r_0\right\}.$$

To prove that the reverse inclusion holds as well, we use the induction hypothesis to deduce that

$$A^{k+1}r_0 = A\left(A^k r_0\right) \in \operatorname{span}\left\{Ap_0, Ap_1, \ldots, Ap_k\right\}$$

Since $Ap_i = \left(r_{i+1} - r_i\right)/\alpha_i$ for $i = 0, 1, \ldots, k$, it follows that

$$A^{k+1}r_0 \in \operatorname{span}\left\{r_0, r_1, \ldots, r_{k+1}\right\}.$$

By combining this expression with the induction hypothesis, we find that

$$\operatorname{span}\left\{r_0, Ar_0, \ldots, A^{k+1}r_0\right\} \subset \operatorname{span}\left\{r_0, r_1, \ldots, r_k, r_{k+1}\right\}.$$

(7b) follows by $r_k \perp \mathcal{V}_{k-1}$.

□

The space $\operatorname{span}\left\{r_0, Ar_0, \ldots, A^k r_0\right\}$ is called *Krylov subspace*. The CG method belongs to a large class of Krylov subspace iterative methods for solving linear algebraic equation.

3. Convergence Analysis

Theorem 3.1. *Let A be SPD and let x_k be the kth iteration in the CG method with an initial guess x_0. Then*

$$\|x^* - x_k\|_A = \inf_{v \in \mathcal{V}_{k-1}} \|x^* - (x_0 + v)\|_A, \tag{8a}$$

$$\|x^* - x_k\|_A = \inf_{p_k \in \mathcal{P}_k, p_k(0)=1} \|p_k(A)\left(x^* - x_0\right)\|_A, \tag{8b}$$

$$\|x^* - x_k\|_A \leq \inf_{p_k \in \mathcal{P}_k, p_k(0)=1} \sup_{\lambda \in \sigma(A)} |p_k(\lambda)| \, \|x^* - x_0\|_A, \tag{8c}$$

where $\mathcal{P}_k$ denotes the set of at most k-degree polynomials and $\sigma(A)$ denotes the set of eigenvalues of A.

Proof. The first identity is from the fact $x^* - x_k = \left(I - \operatorname{Proj}^A_{\mathcal{V}_{k-1}}\right)(x^* - x_0)$. For $v \in \mathcal{V}_{k-1}$, it can be expanded as

$$v = \sum_{i=0}^{k-1} c_i A^i r_0 = \sum_{i=1}^{k} c_{i-1} A^i \left(x_0 - x^*\right).$$

Let $p_k(t) = 1 + \sum_{i=1}^{k} c_{i-1} t^i$. Then

$$x^* - (x_0 + v) = p_k(A)\left(x^* - x_0\right).$$

The identity (8b) then follows from (8a). Since A^i is symmetric in the A-inner product, we have

$$\|p_k(A)\|_A = \rho\left(p_k(A)\right) = \sup_{\lambda \in \sigma(A)} |p_k(\lambda)|$$

which leads to the estimate (8c). □

The polynomial $p_k \in \mathcal{P}_k$ with constraint $p_k(0) = 1$ will be called the *residual polynomial*. Various convergence results of CG method can be obtained by choosing specific residual polynomials.

Theorem 3.2. *If A has only r distinct eigenvalues, then the CG iteration will terminate at the solution in at most r iterations.*

Proof. Suppose that the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r$ take on the r distinct values. We define a polynomial $p_r(t)$ by

$$p_r(t) = \frac{(-1)^r}{\lambda_1 \lambda_2 \cdots \lambda_r} (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_r)$$

and note that $p_r(\lambda_i) = 0$ for $i = 1, 2, \ldots, n$ and $p_r(0) = 1$. $\square$

Remark 3.3. *CG method can be also applied to symmetric and positive semi-definite matrix A. Let $\{\phi_i\}_{i=1}^k$ be the eigenvectors associated to $\lambda_{min}(A) = 0$. Then from Theorem 3.1 , if $b \in \operatorname{range}(A) = \operatorname{span}\{\phi_{k+1}, \phi_{k+2}, \cdots, \phi_n\}$, then the CG method with $x_0 \in \operatorname{range}(A)$ will find a solution $Ax = b$ within $n - k$ iterations.*

CG is invented as a direct method but it is more effective to use as an iterative method. The rate of convergence depends crucially on the distribution of eigenvalues of A and could converge to the solution within certain tolerance in steps $k \ll n$.

Theorem 3.4. *Let x_k be the k-th iteration of the CG method with x_0. Then*

$$\|x^* - x_k\|_A \le 2\left(\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1}\right)^k \|x^* - x_0\|_A$$

Proof. Introduce $T_k(x)$, the Chebyshev polynomial of degree k,

$$T_k(t) = \begin{cases} \cos(k \cdot \arccos t) & \text{if} \quad |t| \le 1, \\ \cosh(k \cdot \operatorname{arccosh} t) & \text{if} \quad |t| \ge 1. \end{cases}$$

To show $T_k(t)$ is indeed a polynomial of x, we can denote by $\theta = \arccos t$ and use

$$(\cos\theta + \mathrm{i}\sin\theta)^k = \left(e^{\mathrm{i}\theta}\right)^k = e^{\mathrm{i}k\theta} = \cos k\theta + \mathrm{i}\sin k\theta.$$

On the left hand side, the real part will contain $(\sin\theta)^{2\ell} = \left(1 - \cos^2\theta\right)^\ell = \left(1 - t^2\right)^\ell$ which is a polynomial of t. For $|t| \ge 1$, verification is similar.

Let $a = \lambda_{\min}(A)$ and $b = \lambda_{\max}(A)$. We use the transformation $t \mapsto \frac{b+a-2t}{b-a}$ to change the interval $[a, b]$ to $[1, -1]$ and can use Chebyshev polynomial to define a residual polynomial

$$p_k(t) = \frac{T_k((b + a - 2t)/(b - a))}{T_k((b + a)/(b - a))}.$$

The denominator is introduced to satisfy the condition $p_k(0) = 1$. For $t \in [a, b]$, the transformed variable

$$\left|\frac{b + a - 2t}{b - a}\right| \le 1.$$

Hence the numerator is $\cos k\theta$ and $|\cos k\theta| \le 1$ which leads to the bound

$$\inf_{p_k \in \mathcal{P}_k, p_k(0)=1} \sup_{\lambda \in \sigma(A)} |p_k(\lambda)| \le \left[T_k\left(\frac{b + a}{b - a}\right)\right]^{-1}.$$

We set

$$\frac{b + a}{b - a} = \cosh\sigma = \frac{\mathrm{e}^\sigma + \mathrm{e}^{-\sigma}}{2}$$

Solving this equation for e^σ, we have

$$\mathrm{e}^\sigma = \frac{\sqrt{\kappa(A)} + 1}{\sqrt{\kappa(A)} - 1}$$

with $\kappa(A) = b/a$. We then obtain

$$T_k\left(\frac{b+a}{b-a}\right) = \cosh(k\sigma) = \frac{\mathrm{e}^{k\sigma} + \mathrm{e}^{-k\sigma}}{2} \geq \frac{1}{2}e^{k\sigma} = \frac{1}{2}\left(\frac{\sqrt{\kappa(A)}+1}{\sqrt{\kappa(A)}-1}\right)^k,$$

which complete the proof. □

The estimate in Theorem 3.4 shows that CG is in general better than the gradient method. Furthermore if the condition number of A is close to one, CG iteration will converge very fast. In some scenario, even if $\kappa(A)$ is large, the iteration will perform well if the majority of eigenvalues are clustered in a few small intervals.

Corollary 3.5. *Assume that $\sigma(A) = \sigma_0(A) \cup \sigma_1(A)$ and l is the number of elements in $\sigma_0(A)$. Then*

$$\|x^* - x_k\|_A \leq 2M\left(\frac{\sqrt{b/a}-1}{\sqrt{b/a}+1}\right)^{k-l} \|x^* - x_0\|_A$$

where

$$a = \min_{\lambda\in\sigma_1(A)} \lambda, b = \max_{\lambda\in\sigma_1(A)} \lambda, \text{ and } M = \max_{\lambda\in\sigma_1(A)} \prod_{\mu\in\sigma_0(A)} |1 - \lambda/\mu|$$

Proof. Take $p_k(t) = \frac{(-1)^l}{\lambda_1\lambda_2\cdots\lambda_l}\left(t-\lambda_1\right)\left(t-\lambda_2\right)\cdots\left(t-\lambda_l\right)\frac{T_{k-l}((b+a-2t)/(b-a))}{T_{k-1}((b+a)/(b-a))}$. □

This result shows that if there are only few (say 2 or 3) small eigenvalues and others are well conditioned (in the sense that the so-called effective condition number b/a is not too large), then after few steps, the convergence rate of CG is governed by the effective condition number b/a.

References

[1] L. Chen. Conjugate gradient methods. *Leture note*, 2020. 1

[2] M. R. Hestenes, E. Stiefel, et al. *Methods of conjugate gradients for solving linear systems*, volume 49. NBS Washington, DC, 1952. 1

[3] J. Nocedal and S. J. Wright. *Numerical optimization*. Springer, 1999. 1