Lecture 10: CS395T Numerical Optimization for Graphics and AI — Conjugate Gradient Methods (Linear)

Qixing Huang
The University of Texas at Austin
huangqx@cs.utexas.edu

1 Disclaimer

This note is adapted from


2 Basic Setting

In this section, we study how to solve the following optimization problem using conjugate gradient method

\[
\text{minimize } x \rightarrow -b^T x + \frac{1}{2} x^T A x,
\]

where \( A \) is assumed to be positive definite. Later we will replace this constraint and also consider non-linear objective functions.

**Conjugate directions.** We say a few directions \( p_i, 1 \leq i \leq n \) are conjugate with respect to \( A \) if

\[
p_i^T A p_j = 0, \quad 1 \leq i \neq j \leq n.
\]

It is easy to see that when \( A = I_n \), then conjugate directions become orthogonal directions. This means conjugate directions are essentially generalizations of orthogonal directions.

Conjugate directions also exist, for example, let the spectral decomposition of \( A \) be

\[
A = U \Sigma U^T.
\]

Then we can let \( p_1, \cdots, p_n \) be the columns of \( U \), i.e., \( U = (p_1, \cdots, p_n) \), and it is easy to check that they are indeed conjugate directions. On the other hand, the set of conjugate directions are not unique.

**Fact 2.1.** Conjugate directions are linearly independent.

**Proof.** Suppose they are linearly dependent. This means there exist non-zero (meaning not all of them are zero) coefficients \( c_i, 1 \leq i \leq n \) such that

\[
\sum_{i=1}^{n} c_i p_i = 0.
\]
It follows that
\[ 0 = \left( \sum_{i=1}^{n} c_i p_i \right)^T A \left( \sum_{i=1}^{n} c_i p_i \right) = \sum_{i=1}^{n} c_i^2 (p_i^T A p_i) > 0, \tag{1} \]
which leads to a contradiction.

The conjugate directions are useful for optimization. Intuitively, in the case \( A = I_n \), we can essentially optimize along each coordinate independently to obtain the optimal solution. This is also true in the general case. In fact, starting from an initial solution \( x_0 \), we can gradually improve the solution by searching along \( p_k \) at each iteration:
\[ x_{k+1} = x_k + \alpha_k p_k, \]
where
\[
\alpha_k = \min_{\alpha} b^T (x_{k-1} + \alpha p_k) - \frac{1}{2} (x_{k-1} + \alpha p_k)^T A (x_{k-1} + \alpha p_k)
= -\frac{1}{2} \alpha^2 (p_k^T A p_k) + \alpha (b - Ax_k)^T p_k
= -\frac{r_k^T p_k}{p_k^T A p_k},
\]
where \( r_k := b - Ax_k \).

**Proposition 1.** The procedure described above converges to the optimal solution \( x^* = -A^{-1}b \) in at most \( n \) steps.

**Proof.** Define
\[ r_k := Ap_k - b. \]
It is equivalent to show that
\[ r_k^T p_i = 0, \quad 1 \leq i < k \leq n. \]
We prove this by induction. When \( k = 1 \), we have
\[
\begin{align*}
r_1^T p_0 &= -(b - Ax_0 - \alpha_0 Ap_0)^T p_1 \\
&= r_0^T p_0 - \alpha_0 p_0^T A p_0 = 0.
\end{align*}
\]
Suppose it is true for \( 1 \leq i < k \leq j \), now let us consider \( k = j + 1 \). First of all, we have
\[
\begin{align*}
r_{j+1}^T p_j &= -(b - Ax_j - \alpha_j Ap_j)^T p_j \\
&= r_j^T p_j - \alpha_j (p_j^T A p_j) = 0.
\end{align*}
\]
When \( i < j + 1 \), we have
\[
\begin{align*}
r_{j+1}^T p_i &= -(b - Ax_j - \alpha_j Ap_j)^T p_i \\
&= -r_j^T p_i + \alpha_j (p_j^T A p_i) = 0,
\end{align*}
\]
which ends the proof.

The conjugate gradient method is based on a genius idea of choosing the conjugate gradient directions. Specifically, let
\[ p_0 = -r_0 = b - Ax_0. \]
The intermediate search directions are given by
\[ p_k = -r_k + \beta_k p_{k-1}, \]
where $\beta_k$ is chosen such that $p_k^T A p_{k-1} = 0$ or in other words

$$
\beta_k = \frac{r_k^T A p_{k-1}}{p_{k-1}^T A p_{k-1}}.
$$

Interesting, we can show that $p_k$ is also conjugate to previous directions. Formally speaking, we summarize the properties in the following theorem:

**Theorem 2.1.** Suppose that the $k$th iterate generated by the conjugate gradient method is not the solution point $x^*$. The following four properties hold:

1. $r_k^T r_i = 0$, for $i = 0, 1, \ldots, k - 1$, (3)
2. $\text{span}\{r_0, r_1, \ldots, r_k\} = \text{span}\{r_0, Ar_0, \ldots, A^k r_0\}$, (4)
3. $\text{span}\{p_0, p_1, \ldots, p_k\} = \text{span}\{r_0, Ar_0, \ldots, A^k r_0\}$, (5)
4. $p_k^T A p_i = 0$, for $i = 0, 1, \ldots, k - 1$. (6)

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

**Proof.** The proof is by induction. The expressions (4) and (5) holds trivially when $k = 0$. In addition, by construction $p_0^T A p_0 = 0$ so (6) also holds when $k = 1$. Now suppose (3)-(6) hold for all $j \leq k$. Now consider $k + 1$. Since $p_0, \ldots, p_k$ are conjugate directions, we conclude that

$$
r_k^T p_i = 0, \quad 0 \leq i \leq k.
$$

We first show that

$$
\alpha_k \neq 0.
$$

Suppose $\alpha_k = 0$, this means

$$
0 = r_k^T A p_k
= r_k^T A (-r_k + \beta_k p_{k-1})
= -r_k^T A r_k + \frac{1}{\alpha_{k-1}} r_k^T (r_{k-1} - r_{k-2})
= -r_k^T A r_k.
$$

Since $A$ is positive semidefinite, it means $r_k = 0$, which leads to a contradiction.

We then consider (4) and (5). Since $r_{k+1} = \alpha_k A p_k + r_k$. As by induction $p_k \in \text{span}\{r_0, Ar_0, \ldots, A^k r_0\}$. It follows that $r_{k+1} \in \text{span}\{r_0, Ar_0, \ldots, A^{k+1} r_0\}$. This means

$$
\text{span}\{r_0, r_1, \ldots, r_{k+1}\} \subset \text{span}\{r_0, Ar_0, \ldots, A^{k+1} r_0\}.
$$

In addition, since $p_{k+1} = -r_{k+1} + \beta_{k+1} p_k$, it follows that

$$
\text{span}\{p_0, p_1, \ldots, p_{k+1}\} \subset \text{span}\{r_0, Ar_0, \ldots, A^{k+1} r_0\}.
$$

To prove the inverse. Let

$$
A^k r_0 = c_0 r_0 + \cdots + c_k r_k.
$$
It follows that
\[
A^{k+1}r_0 \in \text{span}\{Ar_0, \cdots, Ar_k\}
\]
\[
\in \text{span}\{Ar_0, A(\beta_1p_0 - p_1), \cdots, A(\beta_kp_{k-1} - p_k)\}
\]
\[
\in \text{span}\{Ap_0, Ap_1, \cdots, Ap_k\}
\]
\[
\in \text{span}\left\{\frac{(r_1 - r_0)}{\alpha_0}, \frac{(r_2 - r_1)}{\alpha_1}, \cdots, \frac{(r_k - r_{k-1})}{\alpha_k}\right\}
\]
\[
\in \text{span}\{r_0, \cdots, r_k\}
\]
\[
\in \text{span}\{p_0, \beta_1p_0 - p_1, \cdots, \beta_kp_k - p_{k+1}\}
\]
\[
\in \text{span}\{p_0, p_1, \cdots, p_{k+1}\}
\]

Now we show that
\[
r_{k+1}^Tr_i = 0, \quad 0 \leq i \leq k.
\]
In fact,
\[
r_{k+1}^Tr_i = r_{k+1}^T(\beta_i p_{i-1} - p_i) = \beta_i(r_{k+1}^Tp_{i-1}) - r_{k+1}^Tp_i = 0.
\]
Finally, for \(0 \leq i \leq k\),
\[
p_{k+1}^TAp_i = (-r_{k+1} + \beta_kp_k)^TAp_i = -r_{k+1}^TAp_i = -r_{k+1}^T\left(\frac{r_i - r_{i-1}}{\alpha_i}\right) = 0.
\]

Conjugate gradient methods require computing two constants
\[
\beta_k = \frac{r_{k+1}^TAp_k}{p_k^TAp_k}, \quad \alpha_k = -\frac{r_k^TAp_k}{p_k^TAp_k}.
\]
Note that
\[
r_{k}^TAp_k = -r_{k}^T(-r_{k} + \beta_kp_{k-1}) = r_{k}^Tr_k.
\]
Moreover,
\[
\beta_k = \frac{r_{k+1}^TAp_k}{p_k^TAp_k} = \frac{r_{k+1}^TAp_k}{p_k^TAp(-r_k + \beta_kp_{k-1})} = \frac{r_{k+1}^TAp_k}{r_k^TAp_k} = -\frac{r_{k+1}^T(\alpha_kAp_k)}{r_k^T(\alpha_kAp_k)}
\]
\[
= \frac{-r_{k+1}^T(r_k - r_{k-1})}{r_k^T(r_k - r_{k-1})} = \frac{r_{k+1}^Tr_{k+1}}{r_k^Tr_k}.
\]
Finally, we arrive at the standard form of conjugate gradient descent method:

- Given \(x_0\);
- Set \(r_0 = Ax_0 - b, p_0 \leftarrow -r_0, k \leftarrow 0\);
- while \(r_k \neq 0\)
  - \(\alpha_k \leftarrow \frac{r_k^Tr_k}{p_k^TAp_k}\);
  - \(x_{k+1} \leftarrow x_k + \alpha_kp_k\);
  - \(r_{k+1} \leftarrow r_k + \alpha_kAp_k\);
  - \(\beta_{k+1} \leftarrow \frac{r_{k+1}^Tr_{k+1}}{r_k^Tr_k}\);
  - \(p_{k+1} = -r_{k+1} + \beta_{k+1}p_k\);
  - \(k \leftarrow k + 1\);
- end (while)
2.1 Convergence Rate of CG

One thing is analyze the convergence of CG is to utilize the fact that $x_{k+1}$ is the approximation of $x^*$ in $x_0 + \text{span}(r_0, Ar_0, \ldots, A^k r_0)$. Define

$$P_k(A) = \gamma_0 I + \gamma_1 A + \cdots + \gamma_k A^k,$$

where $\gamma_0, \gamma_1, \cdots, \gamma_k$ are coefficients. It turns out $x_{k+1}$ is given by the best degree $k$ polynomial that solves the following optimization problem:

$$\min_{P_k} \|x_0 + P_k(A)r_0 - x^*\|^2_A.$$

Let $P^*_k(A)$ be the best polynomial, we have that

$$x_{k+1} - x^* = x_0 + P^*_k(A)r_0 - x^* = (I + P^*_k(A)A)(x_0 - x^*).$$

Let $v_i, 1 \leq i \leq n$ be the eigen-vectors of $A$. Write

$$x_0 - x^* = \sum_{i=1}^n \psi_i v_i.$$ 

It follows that

$$\|x_{k+1} - x^*\|^2_A = \sum_{i=1}^n \lambda_i (1 + \lambda_i P^*_k(\lambda_i)) \psi_i^2.$$ 

So we have

$$\|x_{k+1} - x^*\|^2_A = \min_{P_k} \sum_{i=1}^n \lambda_i (1 + \lambda_i P_k(\lambda_i)) \psi_i^2.$$ 

It follows that

$$\|x_{k+1} - x^*\|^2_A \leq \min_{P_k} \max_{1 \leq i \leq n} (1 + \lambda_i P_k(\lambda_i))^2 \|x_0 - x^*\|^2_A.$$

By choosing different polynomials, we can have different types convergence bounds. In the literature, people have obtained the following bounds:

$$\|x_{k+1} - x^*\|^2_A \leq \left(\frac{\lambda_{k+1} - \lambda_n}{\lambda_{k+1} + \lambda_n}\right)^2 \left\|x_0 - x^*\right\|^2_A \tag{7}$$

and

$$\|x_{k+1} - x^*\|^2_A \leq \left(\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1}\right)^{k+1} \left\|x_0 - x^*\right\|^2_A \tag{8}.$$