Outline

1 Recap
   - Trust Region: Global Convergence and Enhancements

2 Conjugate Gradient Methods
   - Introduction: Notation, Definitions, Properties
   - A Conjugate Direction Method

3 A Little Bit (More) Theory...
   - $n$-step Convergence for Non-Diagonal $A$; Cheap Residuals
   - Expanding Subspace Minimization

Peter Blomgren, ⟨blomgren.peter@gmail.com⟩  Linear CG, Part #1
Quick Recap: — Global Convergence and Enhancements

We looked at some theorems describing the convergence of our algorithms. We noted that there was a bit of a gap between what is generally true/practical, and what can be proved. (Theoretical limit points vs. numerical stopping criteria.)

Further, we looked at some enhancements including scaling

$$D = \text{diag}(d_1, d_2, \ldots, d_n), \quad d_i > 0, \quad T(\Delta) = \{\bar{p} \in \mathbb{R}^n : \|D\bar{p}\| \leq \Delta\},$$

and the use of non-Euclidean norms — the latter primarily come in handy in the context of constrained optimization.

We now explore an important computational tool, which will help us solve problems of realistic size. — Conjugate Gradient Methods.
For short: “CG” Methods.

- One of the most useful techniques for solving large linear systems of equations $A\tilde{x} = \tilde{b}$. “Linear CG”
- Can be adopted to solve nonlinear optimization problems. “Nonlinear CG” (Our type of problems!)
- Linear CG is an alternative to Gaussian elimination (well suited for large problems).
- Performance of linear CG is strongly tied to the distribution of the eigenvalues of $A$.

First, we explore the Linear CG method...
The linear CG method is an iterative method for solving linear systems of equations:

\[ A\tilde{x} = \tilde{b}, \quad A \in \mathbb{R}^{n \times n}, \quad \tilde{x} \in \mathbb{R}^n, \quad \tilde{b} \in \mathbb{R}^n, \]

where the matrix \( A \) is symmetric positive definite.

Notice/Recall: This problem is equivalent to minimizing \( \Phi(\tilde{x}) \) where

\[ \Phi(\tilde{x}) = \frac{1}{2} \tilde{x}^T A \tilde{x} - \tilde{b}^T \tilde{x} + c, \]

since

\[ \nabla \Phi(\tilde{x}) = A\tilde{x} - \tilde{b} \equiv \tilde{r}(\tilde{x}). \]

We refer to \( \tilde{r}(\tilde{x}) \) as the residual of the linear system. Note that if \( \tilde{x}^* = A^{-1}\tilde{b} \), then \( \tilde{r}(\tilde{x}^*) = 0 \), i.e. the residual is a measure of how close (or far) we are from solving the linear system.
Conjugate Directions

Definition (Conjugate Vector)

A set of nonzero vectors \( \{\vec{p}_0, \vec{p}_1, \ldots, \vec{p}_{n-1}\} \) is said to be conjugate with respect to the symmetric positive definite matrix \( A \) if

\[
\vec{p}_i^T A \vec{p}_j = 0, \quad \forall i \neq j.
\]

Property: Linear Independence of Conjugate Vectors

A set of conjugate vectors \( \{\vec{p}_0, \vec{p}_1, \ldots, \vec{p}_{n-1}\} \) is linearly independent.

Why should we care? — We can minimize \( \Phi(\bar{x}) \) in \( n \) steps by successively minimizing along the directions in a conjugate set...
Given a starting point $\bar{x}_0 \in \mathbb{R}^n$, and a set of conjugate directions $\{\bar{p}_0, \bar{p}_1, \ldots, \bar{p}_{n-1}\}$ we generate a sequence of points $\bar{x}_k \in \mathbb{R}^n$ by setting

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

where $\alpha_k$ is the minimizer of the quadratic function $\varphi(\alpha) = \Phi(\bar{x}_k + \alpha \bar{p}_k)$, i.e. the minimizer of $\Phi(\cdot)$ along the line $\ell(\alpha) = \bar{x}_k + \alpha \bar{p}_k$.

We have already solved this problem — in the context of step-length selection for line search methods, see lecture #6 — so we “know” that the optimizer is given by

$$\alpha_k = -\frac{\bar{r}_k^T \bar{p}_k}{\bar{p}_k^T A \bar{p}_k}, \quad \text{where } \bar{r}_k = \bar{r}(\bar{x}_k).$$
Theorem *(n-step convergence)*

For any $\bar{x}_0 \in \mathbb{R}^n$ the sequence $\{\bar{x}_k\}$ generated by the conjugate direction algorithm converges to the solution $\bar{x}^*$ of the linear system in at most $n$ steps.

The proof indicates how properties of CG are found...

**Proof: Part 1.**
Theorem ($n$-step convergence)

For any $\bar{x}_0 \in \mathbb{R}^n$ the sequence $\{\bar{x}_k\}$ generated by the conjugate direction algorithm converges to the solution $\bar{x}^*$ of the linear system in at most $n$ steps.

The proof indicates how properties of CG are found...


Since the directions $\{\bar{p}_i\}$ are linearly independent, they must span the whole space $\mathbb{R}^n$. Hence, we can write

$$\bar{x}^* - \bar{x}_0 = \sum_{k=0}^{n-1} \sigma_k \bar{p}_k$$

for some choice of scalars $\sigma_k$. We need to establish that $\sigma_k = \alpha_k$. 

Peter Blomgren, (blomgren.peter@gmail.com)
Proof: Part 2.

If we are generating $\bar{x}_k$ by the conjugate direction method, then we have

$$\bar{x}_k = \bar{x}_0 + \alpha_0 \bar{p}_0 + \alpha_1 \bar{p}_1 + \cdots + \alpha_{k-1} \bar{p}_{k-1},$$
Proof: Part 2.

If we are generating $\bar{x}_k$ by the conjugate direction method, then we have

$$\bar{x}_k = \bar{x}_0 + \alpha_0 \bar{p}_0 + \alpha_1 \bar{p}_1 + \cdots + \alpha_{k-1} \bar{p}_{k-1},$$

we multiply this by $\bar{p}_k^T A$

$$\bar{p}_k^T A \bar{x}_k = \bar{p}_k^T A [\bar{x}_0 + \alpha_0 \bar{p}_0 + \alpha_1 \bar{p}_1 + \cdots + \alpha_{k-1} \bar{p}_{k-1}],$$
Proof: Part 2.

If we are generating $\bar{x}_k$ by the conjugate direction method, then we have

$$
\bar{x}_k = \bar{x}_0 + \alpha_0 \bar{p}_0 + \alpha_1 \bar{p}_1 + \cdots + \alpha_{k-1} \bar{p}_{k-1},
$$

we multiply this by $\bar{p}_k^T A$

$$
\bar{p}_k^T A \bar{x}_k = \bar{p}_k^T A [\bar{x}_0 + \alpha_0 \bar{p}_0 + \alpha_1 \bar{p}_1 + \cdots + \alpha_{k-1} \bar{p}_{k-1}],
$$

using the conjugacy property, we see that all but the first term on the right-hand-side are zero:

$$
\bar{p}_k^T A \bar{x}_k = \bar{p}_k^T A \bar{x}_0 \iff \bar{p}_k^T A (\bar{x}_k - \bar{x}_0) = 0.
$$
Proof: Part 2.

If we are generating $\bar{x}_k$ by the conjugate direction method, then we have

$$\bar{x}_k = \bar{x}_0 + \alpha_0 \bar{p}_0 + \alpha_1 \bar{p}_1 + \cdots + \alpha_{k-1} \bar{p}_{k-1},$$

we multiply this by $\bar{p}_k^T A$

$$\bar{p}_k^T A \bar{x}_k = \bar{p}_k^T A [\bar{x}_0 + \alpha_0 \bar{p}_0 + \alpha_1 \bar{p}_1 + \cdots + \alpha_{k-1} \bar{p}_{k-1}],$$

using the conjugacy property, we see that all but the first term on the right-hand-side are zero:

$$\bar{p}_k^T A \bar{x}_k = \bar{p}_k^T A \bar{x}_0 \iff \bar{p}_k^T A (\bar{x}_k - \bar{x}_0) = 0.$$ 

Now we have

$$\bar{p}_k^T A (\bar{x}^* - \bar{x}_0) = \bar{p}_k^T A (\bar{x}^* - \bar{x}_0 - (\bar{x}_k - \bar{x}_0)) = \bar{p}_k^T A (\bar{x}^* - \bar{x}_k) = \bar{p}_k^T (\bar{b} - A \bar{x}_k) = -\bar{p}_k^T \bar{r}_k.$$ 

adds 0
Proof: Part 3.

We have shown
\[ \bar{p}_k^T A(\bar{x}^* - \bar{x}_0) = -\bar{p}_k^T \bar{r}_k. \]

Now, we notice that the right-hand-side is the numerator in \( \alpha_k \):
\[
\alpha_k = \frac{-\bar{p}_k^T \bar{r}_k}{\bar{p}_k^T A \bar{p}_k} \Rightarrow \alpha_k = \frac{\bar{p}_k^T A(\bar{x}^* - \bar{x}_0)}{\bar{p}_k^T A \bar{p}_k}.
\]
Proof: Part 3.

We have shown

\[ \bar{p}_k^T A(\bar{x}^* - \bar{x}_0) = -\bar{p}_k^T \bar{r}_k. \]

Now, we notice that the right-hand-side is the numerator in \( \alpha_k \):

\[ \alpha_k = \frac{-\bar{p}_k^T \bar{r}_k}{\bar{p}_k^T A \bar{p}_k} \Rightarrow \alpha_k = \frac{\bar{p}_k^T A(\bar{x}^* - \bar{x}_0)}{\bar{p}_k^T A \bar{p}_k}. \]

We conclude the proof by showing that \( \sigma_k \) can be expressed in the same manner;
Proof: Part 3.

We have shown

\[ \bar{p}_k^T A (\bar{x}^* - \bar{x}_0) = -\bar{p}_k^T \bar{r}_k. \]

Now, we notice that the right-hand-side is the numerator in \( \alpha_k \):

\[ \alpha_k = \frac{-\bar{p}_k^T \bar{r}_k}{\bar{p}_k^T A \bar{p}_k} \Rightarrow \alpha_k = \frac{\bar{p}_k^T A (\bar{x}^* - \bar{x}_0)}{\bar{p}_k^T A \bar{p}_k}. \]

We conclude the proof by showing that \( \sigma_k \) can be expressed in the same manner; we premultiply the expression for \( (\bar{x}^* - \bar{x}_0) \) by \( \bar{p}_k^T A \) and obtain

\[ \bar{p}_k^T A (\bar{x}^* - \bar{x}_0) = \bar{p}_k^T A \sum_{i=0}^{n-1} \sigma_i \bar{p}_i = \sum_{i=0}^{n-1} \sigma_i \bar{p}_k^T A \bar{p}_i = \sigma_k \bar{p}_k^T A \bar{p}_k. \]

Hence,

\[ \sigma_k = \frac{\bar{p}_k^T A (\bar{x}^* - \bar{x}_0)}{\bar{p}_k^T A \bar{p}_k} \equiv \alpha_k. \]
Most of the proofs regarding CD and CG methods are argued in a similar way — by looking at optimizers and residuals over sub-spaces of $\mathbb{R}^n$ spanned by some subset of a set of conjugate vectors.

**Interpretation:** If the matrix $A$ is diagonal, then the contours of $\Phi(\bar{x})$ are ellipses whose axes are aligned with the coordinate directions. In this case, we can find the minimizer by performing 1D-minimizations along the coordinate directions $\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n$ in turn.
Interpretation (ctd.): When $A$ is not diagonal, the contours are still elliptical, but are no longer aligned with the coordinate axes. Successive minimization along the coordinate directions $\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n$ can **not** guarantee convergence in $n$ (or even a (fixed) finite number of) iterations.
For non-diagonal matrices $A$, the $n$-step convergence can be recovered by transforming the problem.

Let $S \in \mathbb{R}^{n \times n}$ be a matrix with conjugate columns, i.e. if \{\(\vec{p}_0, \vec{p}_1, \ldots, \vec{p}_{n-1}\)\} is a set of conjugate directions (with respect to $A$), then

$$S = \begin{bmatrix} \vec{p}_0 & \vec{p}_1 & \cdots & \vec{p}_{n-1} \end{bmatrix}.$$ 

We introduce a new variable $\hat{x} = S^{-1}x$, and thus get the new quadratic objective which can be minimized in $n$ steps

$$\hat{\Phi}(\hat{x}) = \Phi(S\hat{x}) = \frac{1}{2} \hat{x}^T (S^T AS) \hat{x} - (S^T \vec{b})^T \hat{x}.$$ 

Diagonal

Peter Blomgren, ⟨blomgren.peter@gmail.com⟩
We note that the matrix \((S^T AS)\) is diagonal by the conjugacy property, and that each coordinate direction \(\hat{e}_i\) in \(\hat{x}\)-space corresponds to the direction \(\bar{p}_{i-1}\) in \(\bar{x}\)-space.

When the matrix is diagonal, each coordinate minimization determines one of the components of the solution \(\bar{x}^*\). Hence, after \(k\) iterations, the quadratic has been minimized on the subspace spanned by \(\hat{e}_1, \hat{e}_2, \ldots, \hat{e}_k\).

If we instead minimize along the conjugate directions, then after \(k\) iterations, the quadratic has been minimized on the subspace spanned by \(\bar{p}_0, \bar{p}_1, \ldots, \bar{p}_{k-1}\).
Before we state a fundamental theorem regarding the conjugate direction method, we show the following lemma:

**Lemma**

Given a starting point \( \bar{x}_0 \in \mathbb{R}^n \) and a set of conjugate directions \( \{\bar{p}_0, \bar{p}_1, \ldots, \bar{p}_{n-1}\} \) if we generate the sequence \( \bar{x}_k \in \mathbb{R}^n \) by setting

\[
\bar{x}_{k+1} = \bar{x}_k + \alpha_k \bar{p}_k, \quad \text{where} \quad \alpha_k = -\frac{\bar{r}_k^T \bar{p}_k}{\bar{p}_k^T A \bar{p}_k},
\]

with \( \bar{r}_k = A \bar{x}_k - b \). Then the \((k + 1)\)st residual is given by the following expression

\[
\bar{r}_{k+1} = \bar{r}_k + \alpha_k A \bar{p}_k.
\]
Before we state a fundamental theorem regarding the conjugate direction method, we show the following lemma:

**Lemma**

Given a starting point $\bar{x}_0 \in \mathbb{R}^n$ and a set of conjugate directions $\{\bar{p}_0, \bar{p}_1, \ldots, \bar{p}_{n-1}\}$ if we generate the sequence $\bar{x}_k \in \mathbb{R}^n$ by setting

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

where

$$\alpha_k = -\frac{\bar{r}_k^T \bar{p}_k}{\bar{p}_k^T A \bar{p}_k},$$

with $\bar{r}_k = A\bar{x}_k - \bar{b}$. Then the $(k + 1)$st residual is given by the following expression

$$\bar{r}_{k+1} = \bar{r}_k + \alpha_k A \bar{p}_k.$$

**Proof.**

$$\bar{r}_{k+1} = A\bar{x}_{k+1} - \bar{b} = A(\bar{x}_k + \alpha_k \bar{p}_k) - \bar{b} = \alpha_k A \bar{p}_k + (A\bar{x}_k - \bar{b}) = \alpha_k A \bar{p}_k + \bar{r}_k.$$
Theorem (Expanding Subspace Minimization)

Let $\bar{x}_0 \in \mathbb{R}^n$ be any starting point and suppose that the sequence $\{\bar{x}_k\}$ is generated by

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

where $\alpha_k = -\frac{\bar{r}_k^T \bar{p}_k}{\bar{p}_k^T A \bar{p}_k}$.

Then

$$\bar{r}_k^T \bar{p}_i = 0, \quad \text{for } i = 0, 1, \ldots, k - 1,$$

and $\bar{x}_k$ is the minimizer of $\Phi(\bar{x}) = \frac{1}{2} \bar{x}^T A \bar{x} - \bar{b}^T \bar{x}$ over the set

$$S(\bar{x}_0, k) = \left\{ \bar{x} : \bar{x} = \bar{x}_0 + \text{span}\{\bar{p}_0, \bar{p}_1, \ldots, \bar{p}_{k-1}\} \right\}.$$

First, we show that a point $\tilde{x}$ minimizes $\Phi$ over the set $S(\tilde{x}_0, k)$ if and only if $r(\tilde{x})^T\tilde{p}_i = 0$, $i = 0, 1, \ldots, k - 1$. 

Peter Blomgren, (blomgren.peter@gmail.com)

First, we show that a point $\tilde{x}$ minimizes $\Phi$ over the set $S(\tilde{x}_0, k)$ if and only if $r(\tilde{x})^T\tilde{p}_i = 0, \ i = 0, 1, \ldots, k - 1$.

Let $h(\tilde{\sigma}) = \Phi(\tilde{x}_0 + \sigma_0\tilde{p}_0 + \sigma_1\tilde{p}_1 + \cdots + \sigma_{k-1}\tilde{p}_{k-1})$. Since $h(\tilde{\sigma})$ is a strictly convex quadratic it has a unique minimizer $\tilde{\sigma}^*$ that satisfies

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

First, we show that a point $\tilde{x}$ minimizes $\Phi$ over the set $S(\tilde{x}_0, k)$ if and only if $r(\tilde{x})^T\bar{p}_i = 0$, $i = 0, 1, \ldots, k - 1$.

Let $h(\bar{\sigma}) = \Phi(\tilde{x}_0 + \sigma_0\bar{p}_0 + \sigma_1\bar{p}_1 + \cdots + \sigma_{k-1}\bar{p}_{k-1})$. Since $h(\bar{\sigma})$ is a strictly convex quadratic it has a unique minimizer $\bar{\sigma}^*$ that satisfies

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

By the chain rule, this is equivalent to

$$\nabla \Phi(\tilde{x}_0 + \underbrace{\sigma_0^*\bar{p}_0 + \sigma_1^*\bar{p}_1 + \cdots + \sigma_{k-1}^*\bar{p}_{k-1}}_{\tilde{x}})^T\bar{p}_i = 0, \quad i = 0, 1, \ldots, k - 1$$

First, we show that a point $\tilde{x}$ minimizes $\Phi$ over the set $S(\bar{x}_0, k)$ if and only if $\bar{r}(\tilde{x})^T \bar{p}_i = 0$, $i = 0, 1, \ldots, k - 1$.

Let $h(\bar{\sigma}) = \Phi(\bar{x}_0 + \sigma_0 \bar{p}_0 + \sigma_1 \bar{p}_1 + \cdots + \sigma_{k-1} \bar{p}_{k-1})$. Since $h(\bar{\sigma})$ is a strictly convex quadratic it has a unique minimizer $\bar{\sigma}^*$ that satisfies

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

By the chain rule, this is equivalent to

$$\nabla \Phi(\bar{x}_0 + \sigma_0^* \bar{p}_0 + \sigma_1^* \bar{p}_1 + \cdots + \sigma_{k-1}^* \bar{p}_{k-1})^T \bar{p}_i = 0, \quad i = 0, 1, \ldots, k - 1$$

We recall that $\nabla \Phi(\tilde{x}) = A\tilde{x} - \bar{b} = \bar{r}(\tilde{x})$, thus we have established $\bar{r}(\tilde{x})^T \bar{p}_i = 0 \iff \tilde{x}$ minimizes $\Phi$ over the set $S(\bar{x}_0, k)$. 

Peter Blomgren, ⟨blomgren.peter@gmail.com⟩

Linear CG, Part #1
Proof: Part 2.

We now show that the residuals $\tilde{r}_k$ satisfy $\tilde{r}_k^T \tilde{p}_i = 0, \ i = 0, 1, \ldots, k - 1$. 
Expanding Subspace Minimization: Proof

Proof: Part 2.

We now show that the residuals $\bar{r}_k$ satisfy $\bar{r}_k^T \bar{p}_i = 0$, $i = 0, 1, \ldots, k - 1$.

We use mathematical induction. Since $\alpha_0$ is always the 1D-minimizer, we have $\bar{r}_1^T \bar{p}_0 = 0$, establishing the base case.

From the inductive hypothesis, that $\bar{r}_{k-1}^T \bar{p}_i = 0$, $i = 0, 1, \ldots, k - 2$, we must show that $\bar{r}_k^T \bar{p}_i = 0$, $i = 0, 1, \ldots, k - 1$ in order to complete the proof.
Proof: Part 2.

We now show that the residuals $\bar{r}_k$ satisfy $\bar{r}_k^T \bar{p}_i = 0, \ i = 0, 1, \ldots, k - 1$.

We use mathematical induction. Since $\alpha_0$ is always the 1D-minimizer, we have $\bar{r}_1^T \bar{p}_0 = 0$, establishing the base case.

From the inductive hypothesis, that $\bar{r}_{k-1}^T \bar{p}_i = 0, \ i = 0, 1, \ldots, k - 2$, we must show that $\bar{r}_k^T \bar{p}_i = 0, \ i = 0, 1, \ldots, k - 1$ in order to complete the proof.

From the lemma we have an expression for $\bar{r}_k = \bar{r}_{k-1} + \alpha_{k-1} A \bar{p}_{k-1}$.

First off we have: $\bar{p}_{k-1}^T \bar{r}_k = \bar{p}_{k-1}^T \bar{r}_{k-1} + \alpha_{k-1} \bar{p}_{k-1}^T A \bar{p}_{k-1} = 0$, since, by construction (optimality)

$$\alpha_{k-1} = \frac{-\bar{p}_{k-1}^T \bar{r}_{k-1}}{\bar{p}_{k-1}^T A \bar{p}_{k-1}}$$
Proof: Part 3.

Finally,

\[ \bar{p}_i^T \bar{r}_k = \bar{p}_i^T \bar{r}_{k-1} + \alpha_{k-1} \bar{p}_i^T A \bar{p}_{k-1} = 0, \quad i = 0, 1, \ldots, k - 2 \]

since

\[ \bar{p}_i^T \bar{r}_{k-1} = 0, \quad i = 0, 1, \ldots, k - 2 \]

by the induction hypothesis, and

\[ \bar{p}_i^T A \bar{p}_{k-1} = 0, \quad i = 0, 1, \ldots, k - 2 \]

by conjugacy. This establishes \( \bar{p}_i^T \bar{r}_k = 0, \ i = 0, 1, \ldots, k - 1 \), which completes the proof.
Cliff-Hanger Questions:

- How can we make this useful?
- Given $A$, how do we get a set of conjugate vectors? (They are not for sale at Costco!)
- Even if we have them, why is this scheme any better than Gaussian elimination?
- Where is the gradient?
Index

n-step convergence, 8, 9
conjugate direction method, 7
conjugate vector, 6