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

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

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.

Conjugate Gradient Methods: Introduction

For short: “CG” Methods.

- One of the most useful techniques for solving large linear systems of equations \( A\bar{x} = \bar{b} \). “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

The **linear** CG method is an **iterative method** for solving linear systems of equations:

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

where the matrix \( A \) is **symmetric positive definite**. Notice/Recall: This problem is **equivalent to minimizing** \( \Phi(\bar{x}) \) where

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

since

\[ \nabla \Phi(\bar{x}) = A\bar{x} - \bar{b} \quad \text{def} = \bar{r}(\bar{x}). \]

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

**Why should we care?** — We can minimize \( \Phi(\bar{x}) \) in \( n \) steps by successively minimizing along the directions in a conjugate set...

### Conjugate Directions

**Definition (Conjugate Vector)**

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

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

**Property: Linear Independence of Conjugate Vectors**

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

**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** (Fundamental Building Block).

Since the directions \( \{\bar{p}_j\} \) 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 \( \sigma_k = \alpha_k \).
Conjugate Direction Method ( != CG Method ) 3 of 4

Proof: Part 2.
If we are generating \( \tilde{x}_k \) by the conjugate direction method, then we have
\[
\tilde{x}_k = \tilde{x}_0 + \alpha_0 \tilde{p}_0 + \alpha_1 \tilde{p}_1 + \cdots + \alpha_{k-1} \tilde{p}_{k-1},
\]
we multiply this by \( \tilde{p}_k^T A \)
\[
\tilde{p}_k^T A \tilde{x}_k = \tilde{p}_k^T A [\tilde{x}_0 + \alpha_0 \tilde{p}_0 + \alpha_1 \tilde{p}_1 + \cdots + \alpha_{k-1} \tilde{p}_{k-1}],
\]
using the conjugacy property, we see that all but the first term on the right-hand-side are zero:
\[
\tilde{p}_k^T A \tilde{x}_k = \tilde{p}_k^T A \tilde{x}_0 \iff \tilde{p}_k^T A(\tilde{x}_k - \tilde{x}_0) = 0.
\]
Now we have
\[
\tilde{p}_k^T A(\tilde{x}^* - \tilde{x}_0) = \tilde{p}_k^T A(\tilde{x}^* - \tilde{x}_0 - (\tilde{x}_k - \tilde{x}_0)) = \tilde{p}_k^T A(\tilde{x}^* - \tilde{x}_k) = \tilde{p}_k^T (\tilde{b} - A \tilde{x}_k) = -\tilde{p}_k^T \tilde{r}_k.
\]

Conjugate Direction Method: Comments and Interpretation 1 of 2

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 \( \Phi(\tilde{x}) \) 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 \( \tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_n \) 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 \( \tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_n \) can not guarantee convergence in \( n \) (or even a (fixed) finite number of) iterations.
Recovering $n$-step Convergence for Non-Diagonal $A$  

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 $\{\hat{p}_0, \hat{p}_1, \ldots, \hat{p}_{n-1}\}$ is a set of conjugate directions (with respect to $A$), then

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

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

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

Expanding Subspace Minimization

**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

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

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

Then

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

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

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

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 $h(\tilde{\sigma}) = \Phi(\tilde{x}_0 + \tilde{\sigma}\tilde{p}_0 + \tilde{\sigma}_1\tilde{p}_1 + \cdots + \tilde{\sigma}_{k-1}\tilde{p}_{k-1})$. Since $h(\tilde{\sigma})$ is a strictly convex quadratic it has a unique minimizer $\tilde{\sigma}^*$ that satisfies
\[
\frac{\partial h(\tilde{\sigma}^*)}{\partial \tilde{\sigma}_i} = 0, \quad i = 0, 1, \ldots, k - 1
\]

By the chain rule, this is equivalent to
\[
\nabla \Phi(\tilde{x}_0 + \tilde{\sigma}_0\tilde{p}_0 + \tilde{\sigma}_1\tilde{p}_1 + \cdots + \tilde{\sigma}_{k-1}\tilde{p}_{k-1})^T\tilde{p}_i = 0, \quad i = 0, 1, \ldots, k - 1
\]

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

Proof: Part 2.

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

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

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

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

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

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?
Conjugate Gradient Methods
A Little Bit (More) Theory...

- \( n \)-step convergence, 8
- conjugate direction method, 7
- conjugate vector, 6

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