extended discussion of the conjugate gradient method

Canonical name	ExtendedDiscussionOfTheConjugateGradientMethod
Date of creation	2013-03-22 17:28:24
Last modified on	2013-03-22 17:28:24
Owner	ehremo (15714)
Last modified by	ehremo (15714)
Numerical id	13
Author	ehremo (15714)
Entry type	Topic
Classification	msc 15A06
Classification	msc 90C20

Suppose we wish to solve the system

$$A\mathbf{x} = \mathbf{b} \tag{1}$$

where A is a symmetric positive definite matrix. If we define the function

$$\Phi(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{x}^T\mathbf{b} \tag{2}$$

we realise that solving (**??**) is equivalent to minimizing Φ. This is because if $\mathbf{x}$ is a minimum of Φ then we must have

$$\nabla\Phi(\mathbf{x}) = A\mathbf{x} - \mathbf{b} = \mathbf{0} \tag{3}$$

These considerations give rise to the *steepest descent algorithm*. Given an approximation $\tilde{\mathbf{x}}$ to the solution $\mathbf{x}$, the idea is to improve the approximation by moving in the direction in which Φ decreases most rapidly. This direction is given by the gradient of Φ in $\tilde{\mathbf{x}}$. Therefore we formulate our algorithm as follows

> Given an initial approximation $\mathbf{x}^{(0)}$, for $k \in 1...N$,
>
> $$\mathbf{x}^{(k)} = \mathbf{x}^{(k-1)} + \alpha_k \mathbf{r}^{(k)}$$
>
> where $\mathbf{r}^{(k)} = \mathbf{b} - A\mathbf{x}^{(k-1)} = -\nabla\Phi(\mathbf{x}^{(k-1)})$ and α_k is a scalar to be determined.

$\mathbf{r}^{(k)}$ is traditionally called the *residual vector*. We wish to choose α_k in such a way that we reduce Φ as much as possibile in each iteration, in other words, we wish to minimize the function $\phi(\alpha_k) = \Phi(\mathbf{x}^{(k-1)} + \alpha_k\mathbf{r}^{(k)})$ with respect to α_k

$$\begin{aligned}\phi'(\alpha_k) &= \nabla\Phi(\mathbf{x}^{(k-1)} + \alpha_k\mathbf{r}^{(k)})^T\mathbf{r}^{(k)} \\ &= [A\mathbf{x}^{(k)} - \mathbf{b} + \alpha_k A\mathbf{r}^{(k)}]^T\mathbf{r}^{(k)} \\ &= [-\mathbf{r}^{(k)} + \alpha_k A\mathbf{r}^{(k)}]^T\mathbf{r}^{(k)} \\ &= 0\end{aligned} \quad \Longleftrightarrow \alpha_k = \frac{\mathbf{r}^{(k)T}\mathbf{r}^{(k)}}{\mathbf{r}^{(k)T}A\mathbf{r}^{(k)}}$$

It's possible to demonstrate that the steepest descent algorithm described above converges to the solution $\mathbf{x}$, in an infinite time. The conjugate gradient method improves on this by finding the exact solution after only n iterations. Let's see how we can achieve this.

We say that $\tilde{\mathbf{x}}$ is *optimal* with respect to a direction $\mathbf{d}$ if $\lambda = 0$ is a local minimum for the function $\Phi(\tilde{\mathbf{x}} + \lambda\mathbf{d})$.

In the steepest descent algorithm, $\mathbf{x}^{(k)}$ is optimal with respect to $\mathbf{r}^{(k)}$, but in general it is not optimal with respect to $\mathbf{r}^{(0)}, ..., \mathbf{r}^{(k-1)}$. If we could modify the algorithm such that the optimality with respect to the search directions is preserved we might hope that that $\mathbf{x}^{(n)}$ is optimal with respect to n linearly independent directions, at which point we would have found the exact solution.

Let's make the following modification

$$\mathbf{p}^{(k)} = \begin{cases} \mathbf{r}^{(1)} & \text{if } k = 1 \\ \mathbf{r}^{(k)} - \beta_k \mathbf{p}^{(k-1)} & \text{if } k > 1 \end{cases} \tag{4}$$

$$\mathbf{x}^{(k)} = \mathbf{x}^{(k-1)} + \alpha_k \mathbf{p}^{(k)} \tag{5}$$

where α_k and β_k are scalar multipliers to be determined

We choose α_k in the same way as before, i.e. such that $\mathbf{x}^{(k)}$ is optimal with respect to $\mathbf{p}^{(k)}$

$$\alpha_k = \frac{\mathbf{r}^{(k)T}\mathbf{p}^{(k)}}{\mathbf{p}^{(k)T}A\mathbf{p}^{(k)}} \tag{6}$$

Now we wish to choose β_k such that $\mathbf{x}^{(k)}$, is also optimal with respect to $\mathbf{p}^{(k-1)}$. We require that

$$\left.\frac{\partial \Phi(\mathbf{x}^{(k)} + \lambda \mathbf{p}^{(k-1)})}{\partial \lambda}\right|_{\lambda=0} = [A\mathbf{x}^{(k)} - \mathbf{b}]^T \mathbf{p}^{(k-1)} = 0 \tag{7}$$

Since $\mathbf{x}^{(k)} = \mathbf{x}^{(k-1)} + \alpha_k \mathbf{p}^{(k)}$, and assuming that $\mathbf{x}^{(k-1)}$ is optimal with respect to $\mathbf{p}^{(k-1)}$ (i.e. that $[A\mathbf{x}^{(k-1)} - \mathbf{b}]^T\mathbf{p}^{(k-1)} = 0$) we can rewrite this condition as

$$[A(\mathbf{x}^{(k-1)} + \alpha_k \mathbf{p}^{(k)}) - \mathbf{b}]^T \mathbf{p}^{(k-1)} = \alpha_k [\mathbf{r}^{(k)} - \beta_k \mathbf{p}^{(k-1)}]^T A \mathbf{p}^{(k-1)} = 0 \tag{8}$$

and therefore we obtain the required value for β_k,

$$\beta_k = \frac{\mathbf{r}^{(k)T}A\mathbf{p}^{(k-1)}}{\mathbf{p}^{(k-1)T}A\mathbf{p}^{(k-1)}} \tag{9}$$

Now we want to show that $\mathbf{x}^{(k)}$ is also optimal with respect to $\mathbf{p}^{(1)}, ..., \mathbf{p}^{(k-2)}$, i.e. that

$$[A\mathbf{x}^{(k)} - \mathbf{b}]^T \mathbf{p}^{(j)} = \mathbf{r}^{(k+1)T}\mathbf{p}^{(j)} = 0 \qquad \forall\, j \in 1...k-2 \tag{10}$$

We do this by strong induction on k, assuming that for all $\ell \in 1...k-1, j \in 1...\ell$, $\mathbf{x}^{(\ell)}$ is optimal with respect to $\mathbf{p}^{(j)}$ or equivalently that

$$[A\mathbf{x}^{(\ell)} - \mathbf{b}]^T \mathbf{p}^{(j)} = \mathbf{r}^{(\ell+1)T}\mathbf{p}^{(j)} = 0 \tag{11}$$

Noticing that $\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} + \alpha_k A\mathbf{p}^{(k)}$, we want to show

$$\mathbf{r}^{(k)T}\mathbf{p}^{(j)} + \alpha_k \mathbf{p}^{(k)T} A\mathbf{p}^{(j)} = 0 \tag{12}$$

and therefore since $\mathbf{r}^{(k)T}\mathbf{p}^{(j)} = 0$ by inductive hypothesis, it suffices to prove that

$$\mathbf{p}^{(k)T} A\mathbf{p}^{(j)} = 0 \qquad \forall\ j \in 1...k-2 \tag{13}$$

But, again by the definition of $\mathbf{p}^{(k)}$, this is equivalent to proving

$$\mathbf{r}^{(k)T} A\mathbf{p}^{(j)} - \beta_k \mathbf{p}^{(k-1)T} A\mathbf{p}^{(j)} = 0 \tag{14}$$

and since $\mathbf{p}^{(k-1)T} A\mathbf{p}^{(j)} = 0$ by inductive hypothesis, all we need to prove is that

$$\mathbf{r}^{(k)T} A\mathbf{p}^{(j)} = 0 \tag{15}$$

To proceed we require the following lemma.

Let $V_\ell = \text{span}\{\ A^{\ell-1}\mathbf{r}^{(1)}, ..., A\mathbf{r}^{(1)}, \mathbf{r}^{(1)}\ \}$.

- $\mathbf{r}^{(\ell)}, \mathbf{p}^{(\ell)} \in V_\ell$
- If $\mathbf{y} \in V_\ell$ then $\mathbf{r}^{(k)} \perp \mathbf{y}$ for all $k > \ell$.

Since $\mathbf{p}^{(j)} \in V_j$, it follows that $A\mathbf{p}^{(j)} \in V_{j+1}$, and since $j + 1 < k$,

$$\mathbf{r}^{(k)T}(A\mathbf{p}^{(j)}) = 0 \tag{16}$$

To finish let's prove the lemma. The first point is by induction. The base case $\ell = 1$ holds. Assuming that $\mathbf{r}^{(\ell-1)}, \mathbf{p}^{(\ell-1)} \in V_{\ell-1}$, we have that

$$\mathbf{r}^{(\ell)} = \mathbf{r}^{(\ell-1)} + \alpha_{\ell-1} A\mathbf{p}^{(\ell-1)} \in V_\ell \qquad \mathbf{p}^{(\ell)} = \mathbf{r}^{(\ell)} - \beta_\ell \mathbf{p}^{(\ell-1)} \in V_\ell \tag{17}$$

For the second point we need an alternative characterization of V_ℓ. Since $\mathbf{p}^{(1)}, ..., \mathbf{p}^{(\ell)} \in V_\ell$,

$$\text{span}\{\ \mathbf{p}^{(1)}, ..., \mathbf{p}^{(\ell)}\ \} \subseteq V_\ell \tag{18}$$

By (**??**), we have that if for some $s \in 1...\ell$

$$\mathbf{p}^{(s)} = \lambda_{s-1}\mathbf{p}^{(s-1)} + \cdots + \lambda_1 \mathbf{p}^{(1)} \tag{19}$$

then $\mathbf{p}^{(s)T}\mathbf{r}^{(s)} = 0$, but we know that

$$\mathbf{p}^{(s)T}\mathbf{r}^{(s)} = [\mathbf{r}^{(s)} - \beta_s \mathbf{p}^{(s-1)}]^T \mathbf{r}^{(s)} = \mathbf{r}^{(s)T}\mathbf{r}^{(s)} \tag{20}$$

but were this zero, we'd have $\mathbf{r}^{(s)} = \mathbf{0}$ and we would have solved the original problem. Thus we conclude that $\mathbf{p}^{(s)} \notin \text{span}\{\ \mathbf{p}^{(1)}, ..., \mathbf{p}^{(s-1)}\ \}$, so the vectors $\mathbf{p}^{(1)}, ..., \mathbf{p}^{(s)}$ are linearly independent. It follows that $\dim \text{span}\{\ \mathbf{p}^{(1)}, ..., \mathbf{p}^{(\ell)}\ \} = \ell$, which on the other hand is the maximum possible dimension of V_ℓ and thus we must have

$$V_\ell = \text{span}\{\ \mathbf{p}^{(1)}, ..., \mathbf{p}^{(\ell)}\ \} \tag{21}$$

which is the alternative characterization we were looking for. Now we have, again by (**??**), that if $\mathbf{y} \in V_\ell$, and $\ell < k$, then

$$\mathbf{r}^{(k)T}\mathbf{y} = 0 \tag{22}$$

thus the second point is proven.