Dr J Lellmann

Mathematical Tripos Part II: Michaelmas Term 2014

Numerical Analysis – Lecture 19

Approach 4.20 (Minimization of quadratic function) Let us assume for the time being that A is symmetric and positive definite. A different approach to constructing good iterative methods for solving systems of linear equations $A\boldsymbol{x} = \boldsymbol{b}$ is based on succesive minimization of the quadratic function

$$F_1(\boldsymbol{x}^{(k)}) := \|\boldsymbol{x}^{(k)} - \boldsymbol{x}^*\|_A^2 = \|\boldsymbol{e}^{(k)}\|_A^2 \,, \tag{4.4}$$

where $\|\boldsymbol{y}\|_A := \sqrt{\boldsymbol{y}^T A \boldsymbol{y}}$ is a Euclidean-type distance (with positive definite A), and the minimizer is clearly the exact solution. An equivalent approach is to minimize the quadratic function

$$F(\boldsymbol{x}) = \tfrac{1}{2}\boldsymbol{x}^T A \boldsymbol{x} - \boldsymbol{x}^T \boldsymbol{b} \,, \tag{4.5}$$

which attains its minimum when $\nabla F(\boldsymbol{x}) = A\boldsymbol{x} - \boldsymbol{b} = 0$, and which does not involve the unknown $\boldsymbol{x}^*$. (It is easy to check that $F(\boldsymbol{x}) = \frac{1}{2}F_1(\boldsymbol{x}) - \frac{1}{2}C$, where $C = \boldsymbol{x}^{*T} A \boldsymbol{x}^*$ is a constant independent of k, hence equivalence.) So, we choose an iterative method that provides the condition $F(\boldsymbol{x}^{(k+1)}) < F(\boldsymbol{x}^{(k)})$. For example, both Jacobi and Gauss–Seidel methods do. We, however, strengthen this descent condition a bit, and turn to the following algorithm.

(a) We pick any starting vector $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$; (b) for any k, stop if the *residual* $\boldsymbol{g}^{(k)} = A\boldsymbol{x}^{(k)} - \boldsymbol{b}$ is acceptably small; (c) otherwise, a *search direction* $\boldsymbol{d}^{(k)}$ is generated that satisfies $[\mathrm{d}F(\boldsymbol{x}^{(k)} + \omega \boldsymbol{d}^{(k)})/\mathrm{d}\omega]_{\omega=0} < 0$; (d) finally, the value of $\omega > 0$ that minimizes $F(\boldsymbol{x}^{(k)} + \omega \boldsymbol{d}^{(k)})$ is calculated (we call it $\omega^{(k)}$), and the k-th iteration sets

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \omega^{(k)} \boldsymbol{d}^{(k)}. \tag{4.6}$$

The definition (4.5) implies the identity

$$F(\boldsymbol{x}^{(k)} + \omega \boldsymbol{d}^{(k)}) = F(\boldsymbol{x}^{(k)}) + \omega \boldsymbol{d}^{(k)T} \boldsymbol{g}^{(k)} + \tfrac{1}{2}\omega^2 \boldsymbol{d}^{(k)T} A \boldsymbol{d}^{(k)}, \qquad \omega \in \mathbb{R}, \tag{4.7}$$

where $\boldsymbol{g}^{(k)} = \nabla F(\boldsymbol{x}^{(k)}) = A\boldsymbol{x}^{(k)} - \boldsymbol{b}$. So, the search direction has to satisfy $\boldsymbol{d}^{(k)T}\boldsymbol{g}^{(k)} < 0$, which is possible, because termination occurs when $\boldsymbol{g}^{(k)}$ is zero, and $\omega^{(k)}$ that minimizes the expression (4.7) has the value

$$\omega^{(k)} = -\frac{\boldsymbol{d}^{(k)T}\boldsymbol{g}^{(k)}}{\boldsymbol{d}^{(k)T} A \boldsymbol{d}^{(k)}} \,. \tag{4.8}$$

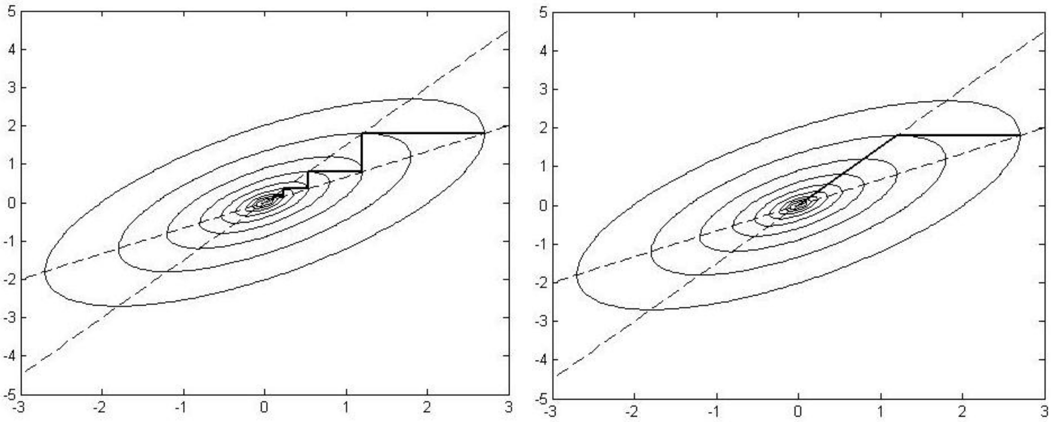
Multiplying both parts of (4.6) with A and subtracting $\boldsymbol{b}$ we see that successive residuals are connected by the rule $\boldsymbol{g}^{(k+1)} = \boldsymbol{g}^{(k)} + \omega^{(k)} A \boldsymbol{d}^{(k)}$, and with the value $\omega^{(k)}$ given above we have the orthogonality condition

$$\boldsymbol{g}^{(k+1)} \;=\; \boldsymbol{g}^{(k)} + \omega^{(k)} A \boldsymbol{d}^{(k)} \perp \boldsymbol{d}^{(k)} \,,$$

Method 4.21 (The steepest descent method) This method makes the choice $\boldsymbol{d}^{(k)} = -\boldsymbol{g}^{(k)}$ for every k, the reason being that, locally, the gradient of a function F shows the direction of the sharpest decay of $F(\boldsymbol{x})$ at $\boldsymbol{x} = \boldsymbol{x}^{(k)}$. Thus, the iterations have the form

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \omega^{(k)}(A\boldsymbol{x}^{(k)} - \boldsymbol{b}), \qquad k \ge 0 \,.$$

It can be proved that, if the number of iterations is infinite, then the sequence $\boldsymbol{x}^{(k)}$, converges to the solution of the system $A\boldsymbol{x} = \boldsymbol{b}$ as required, but usually the speed of convergence is rather slow. The reason is that the iteration decreases the value of $F(\boldsymbol{x}^{(k+1)})$ locally, relatively to $F(\boldsymbol{x}^{(k)})$, but the global decrease, with respect to $F(\boldsymbol{x}^{(0)})$, is often not that large. The use of *conjugate directions* provides a method with a global minimization property.



(a) Worst case scenario of steepest descent

(b) Conjugate gradient method applied to the same problem as in (a)

Figure 1: Courtesy of Anita Briginshaw.

Definition 4.22 (Conjugate directions) The vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ are *conjugate* with respect to the symmetric and positive definite matrix A if they are nonzero and A-orthogonal: $\boldsymbol{u}^T A \boldsymbol{v} = 0$.

The importance of conjugacy to Approach 4.20 depends on the fact that, for the conjugate directions $(\boldsymbol{d}^{(i)})$, the value of $F_1(\boldsymbol{x}^{(k+1)})$ obtained through step-by-step minimization coincides with the minimum of $F_1(\boldsymbol{y})$ taken over all $\boldsymbol{y} = \boldsymbol{x}^{(0)} + \sum_{i=0}^{k} a_i \boldsymbol{d}^{(i)}$ simultaneously, namely

$$\arg \min_{a_0,\ldots,a_k} F_1(\boldsymbol{y}) = \boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(0)} + \sum_{i=0}^{k} \omega^{(i)} \boldsymbol{d}^{(i)} .$$

So, provided that a sequence of conjugate directions $\boldsymbol{d}^{(i)}$ is at hands, we have an iterative procedure with good approximation properties. The algorithm that follows constructs such $\boldsymbol{d}^{(i)}$ by A-orthogonalization of the sequence $(A^i \boldsymbol{g}^{(0)})$. It is of the form described in the second paragraph of Approach 4.20.

Algorithm 4.23 (The conjugate gradient method) Here it is.

(A) For any initial vector $\boldsymbol{x}^{(0)}$, set $\boldsymbol{d}^{(0)} = -\boldsymbol{g}^{(0)} = -(A\boldsymbol{x}^{(0)} - \boldsymbol{b})$;

(B) For $k \geq 0$, calculate $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \omega^{(k)} \boldsymbol{d}^{(k)}$ and the residual

$$\boldsymbol{g}^{(k+1)} = \boldsymbol{g}^{(k)} + \omega^{(k)} A \boldsymbol{d}^{(k)}, \quad \text{with} \quad \omega^{(k)} = \{\, \boldsymbol{g}^{(k+1)} \perp \boldsymbol{d}^{(k)} \} = -\frac{\boldsymbol{d}^{(k)T} \boldsymbol{g}^{(k)}}{\boldsymbol{d}^{(k)T} A \boldsymbol{d}^{(k)}}, \quad k \geq 0. \tag{4.9}$$

(C) For the same k, the next search direction is the vector

$$\boldsymbol{d}^{(k+1)} = -\boldsymbol{g}^{(k+1)} + \beta^{(k)} \boldsymbol{d}^{(k)}, \quad \text{with} \quad \beta^{(k)} = \{\, \boldsymbol{d}^{(k+1)} \perp A \boldsymbol{d}^{(k)} \} = \frac{\boldsymbol{g}^{(k+1)T} A \boldsymbol{d}^{(k)}}{\boldsymbol{d}^{(k)T} A \boldsymbol{d}^{(k)}}, \quad k \geq 0. \tag{4.10}$$

Remark 4.24 *It is possible to lift the restrictive condition on A being symmetric and positive definite by a simple trick. Suppose we want to solve $Bx = c$, where $B \in \mathbb{R}^{n \times n}$ is nonsingular. We can convert the above system to the symmetric and positive definite setting by defining $A = B^T B$, $b = B^T c$ and solving $Ax = b$ with the conjugate gradient algorithm 4.23.*