Assignment 7: Iterative methods

ACM 106a: Introductory Methods of Computational Mathematics (Fall 2013)

Due date: Tuesday, December 3, 2013

1 Gauss-Seidel as convex optimization

In this exercise, we will prove the following:

> *Suppose that the matrix $A \in \mathbf{R}^{n\times n}$ is symmetric positive definite. Then the Gauss-Seidel method applied to the linear system $Ax = b$ converges to the solution $\bar{x} = A^{-1}b$.*

The proof relies on characterizing the Gauss-Seidel method as a *coordinate-descent method* for a particular optimization problem and consists of the following three steps:

(a) Show that if $A \in \mathbf{R}^{n\times n}$ is symmetric positive definite then the solution $\bar{x} = A^{-1}b$ is the unique minimizer of the unconstrained convex program

$$\min_{x\in\mathbf{R}^n} f(x) := \min_{x\in\mathbf{R}^n} \frac{1}{2}x^T Ax - b^T x.$$

Hint: A *convex program* is an optimization problem of the form

$$\min_{x\in\mathbf{R}^n} \{f(x) : x \in C\}$$

where $f : \mathbf{R}^n \to \mathbf{R}$ is a convex function and C is a convex set. In our problem, we have $C = \mathbf{R}^n$. Convex programs have many attractive properties which make them useful for modeling real-world phenomena. Perhaps the most useful is the fact that any stationary point of the problem is a global minimizer. For unconstrained problems (like ours) this means that x is a global minimizer if and only if $\nabla f(x) = 0$.

Hint: To prove uniqueness, it might be useful to consider the value of $f(\bar{x} + d)$ for arbitrary $d \in \mathbf{R}^n$. For example, you could suppose that we have $f(\bar{x} + d) = f(\bar{x})$ for some d and obtain a contradiction using the fact that A is symmetric positive definite.

(b) A *coordinate-descent* algorithm consists, broadly speaking, of an update step accomplished by a perturbation of a single component of the current iterate resulting in a decrease in the objective value. That is, we obtain x^{new} from x by the update formula

$$x^{new} = x + \alpha e_i$$

where e_i is the ith standard basis vector, for some $i \in \{1, 2, \dots, n\}$, and α is a scalar chosen so that $f(x^{new}) \leq f(x)$.

Suppose that we perform a cyclic update in our coordinate-descent algorithm: we first update in direction e_1, then in direction e_2 in the next step, and so on until we update in direction e_n, after which we update in direction e_1 and repeat the cycle. Show that if we greedily choose the step length to maximize decrease in objective value, i.e.,

$$\alpha_i = \arg\min_{\alpha} f(x + \alpha e_i),$$

during each step of coordinate descent then each cycle (from $i = 1$ to $i = n$) is equivalent to one iteration of the Gauss-Seidel method.

Hint: It suffices to prove that we update x_i^{new} by the formula

$$x_i^{new} = \frac{1}{A_{ii}} \left(b_i - \sum_{j \neq i} A_{ij} x_j \right)$$

during the ith step of a cycle of coordinate descent cycle, and confirm that this same operation is used to update x_i in Gauss-Seidel. You can do so by computing an explicit formula for α_i.

(c) Finally, suppose that $x = \bar{x} + e$ is an approximation of $\bar{x} = A^{-1}b$. Show the A-norm defined by

$$\|e\|_A = \sqrt{e^T A e}$$

of the error e satisfies

$$|f(x) - f(\bar{x})| = \frac{1}{2}\|e\|_A^2.$$

Completing the proof: Parts (a), (b), and (c) in tandem prove the theorem. Indeed, Parts (a) and (b) establish that the function value of f decreases during each Gauss-Seidel step (because it decreases during each cycle of coordinate descent unless we have already converged to the solution) and that this value of f is bounded below by the optimal value $f(\bar{x})$. Because these function values form a monotonically decreasing sequence with infimum $f(\bar{x})$, Part (c) implies that the sequence of errors converge to 0 (with respect to the A-norm) or, equivalently, that the sequence of iterates $\{x^{(k)}\}$ generated by Gauss-Seidel converges to $\bar{x}$.

2 Early convergence of Arnoldi iteration

Suppose that at the nth step of Arnoldi iteration (applied to the matrix $A \in \mathbf{C}^{m \times m}$ with vector $b \in \mathbf{C}^m$) we obtain $H_{n+1,n} = 0$ in the recurrence relation

$$AQ_n = Q_{n+1}\tilde{H}_{n+1}.$$

The following exercises will establish that we can stop the Arnoldi iteration after this step; that is, we've found a basis for all Krylov subspaces of A generated by b, and the maximal such subspace contains our desired matrix equation solutions.

(a) Show that we have $AQ_n = Q_n H_n$ in this case.

(b) Using Part (a) show that $\mathcal{K}_n$ is an invariant subspace of A, i.e., $A\mathcal{K}_n \subseteq \mathcal{K}_n$.

Hint: Let x be a linear combination of the vectors $\{q_1, q_2, \ldots, q_n\}$. You can use the identity $AQ_n = Q_n H_n$ to show that Ax is also a linear combination of the vectors $\{q_1, q_2, \ldots, q_n\}$.

(c) Using Part (b), show that we have $\mathcal{K}_n = \mathcal{K}_{n+1} = \cdots$.

Hint: We know that $\mathcal{K}_n \subseteq \mathcal{K}_{n+j}$ for all $j > 0$. To finish the proof, we just have to show that any vector spanned by $\{b, Ab, \ldots, A^{n+j-1}b\}$ must also belong to $\mathcal{K}_n$ for all $j > 0$.

(d) Suppose that λ is an eigenvalue of H_n. Show that λ is also an eigenvalue of A. If v is the corresponding eigenvector of H, what is the corresponding eigenvector of A?

(e) For any nonsingular matrix $A \in \mathbf{C}^{m \times m}$, it is known that we may decompose its inverse as the matrix polynomial

$$A^{-1} = \sum_{k=0}^{m-1} c_k A^k$$

for some scalars $c_0, c_1, \ldots, c_{m-1}$. (This fact is a consequence of the *Cayley-Hamilton Theorem*, which states that the characteristic polynomial of A

$$p(\lambda) = \det(A - \lambda I) = \lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_1\lambda + a_0$$

must satisfy

$$p(A) = A^m + a_{m-1}A^{m-1} + \cdots + a_1 A + a_0 I = 0.)$$

Using this fact, show that the solution of $Ax = b$ belongs to $\mathcal{K}_n$.

3 Practical convergence of Gauss-Seidel and Jacobi

Consider the linear system $Ax = b$ defined by

$$A = \begin{pmatrix} 3 & -5 & 2 \\ 5 & 4 & 3 \\ 2 & 5 & 3 \end{pmatrix}, \qquad b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

(a) Compute the spectral radii of the update matrices R_J and R_{GS} used in the Jacobi and Gauss-Seidel methods. What does this suggest about their usefulness as algorithms for solving $Ax = b$?

The next few parts empirically verify your predictions from Part (a).

(b) Write Matlab code implementing each of the Jacobi and Gauss-Seidel methods in each of the following ways:

(i) by updating each component $x_i^{(k+1)}$ sequentially using a for loop during each iteration, and

(ii) using matrix-vector multiplication updates of the form

$$x^{(k+1)} = Rx^{(k)} + c.$$

Your code should express the update as matrix-vector multiplication using analytic expressions for R_J and R_{GS}, i.e., you should not perform any required inversions or matrix system solves by calling "\".

You should have four (4) Matlab functions. Your functions should take as input A, b, initial iterate $x^{(0)}$, stopping error tolerance *tol*, and maximum number of iterations. That is your code should, starting from $x^{(0)}$, run Jacobi or Gauss-Seidel to solve $Ax = b$ (for arbitrary A and b) until either the relative residual error $\|r^{(k)}\|_2/\|b\|_2$ is less than *tol* (the algorithm converged) or your maximum number of iterations has been exceeded (the algorithm diverges).

(c) Run each of your codes to solve the linear system $Ax = b$ given above starting from the initial iterate $x^{(0)} = (0, 0, 0)$, stopping tolerance $tol = 10^{-3}$, and maximum number of iterations 100. For each trial record the relative error between your solution at termination and the true solution $A^{-1}b$ (obtained using $x = A\backslash b$ in Matlab), the number of iterates required for convergence, and the total time required by each algorithm (as recorded using "tic/toc"). If the time required to solve the problem is too small to distinquish between the algorithms, repeat the experiment many times and compare the aggregate times.

Does your observed phenomena match that predicted in Part (a)? Which algorithms are fastest, those of Type (i) or those of Type (ii)?

(d) Repeat the experiment for a few randomly chosen values of b and $x^{(0)}$. Does the observed convergence phenomena change? Does this coincide with that predicted by theory?

Submission Instructions:

- Assignments are due at the **start** of class (1pm) on the due date.
- Write your name and ID# clearly on all pages, and underline your last name.
- **Matlab files:** please submit a single zip/rar/etc file with file name in the format **Lastname_Firstname_ID#_A7** to **homework.acm106a@gmail.com** that decompresses to a single folder of the same name containing the following:
  - A single thoroughly commented Matlab script file containing all commands used for each assigned programming/simulation problem. File names should have format **Lastname_Firstname_ID#_A7_P#** :

    e.g. **Ames_Brendan_12345678_A7_P4** for problem 4.
  - All Matlab functions used to perform any assigned programming/simulation problems with appropriate file names. Please add a comment to the beginning of each file with the format **Lastname_Firstname_ID#_A7_P#**
  - A diary file of your session for each programming/simulation problem with file name in the format **Lastname_Firstname_ID#_A7_P#_DIARY.txt**. Please also submit a hard copy of the diary and any relevant derivations, pseudocode, etc. you may want considered for partial credit with your submitted solution sets at the beginning of class.