DN2222
Applied Numerical Methods
- part 2:
Numerical Linear Algebra

Lecture 8
Wrap up

2012-12-05

Aim of the course

- The basic course gives a very powerful toolbox, managing many problems.
- But when the problems get BIG, then the "old methods" do not cope. Either they run out-of-memory, take too long time, or just returns rubbish.
- In this course we try to see/explain why some of these things occur and how to avoid them.

This course is looking at two types of problems:

1) Solving Ax=b (when A is $n \times n$ and $m \times n, m > n$)

2) Finding the eigenvalues (or singular values) of A.

- In basic course we learnt to solve Ax=b using Gaussian elimination (with partial pivoting).
- When the size of the problem becomes big, we get into trouble, either from the problem taking too long time to solve OR the obtained solution is not trustworthy.
- The latter problem comes from the fact that the com-

puter is making a round-off error in each calculation.

* We must understand how large the round-off error is and how much effect can come out of this cumulating round-off.

- **Q2.** When is a floating point number normalized?

- **Q4.** What is the overflow threshhold? What is its approximate value in IEEE double precision?

- **Q7.** How do you bound the rounding error when computing a sum of three numbers, $s = a+b+c$?

- **Q10.** What are well-conditioned problems?
- **Q11.** Why is well-conditioning necessary for obtaining reliable results?
- **Q21.** Define the condition number of a matrix.
- **Qxx.** What is meant by partial pivoting of a matrix.
- **Qxx.** What is meant by scaling a matrix.
- **Q9.** What are well-posed and ill-posed problems?

- **Q12.** What are forward and backward stability?
- The total round-off error is also shrunk if the amount of operations needed is lowered.
- The order which we number the nodes/variables is important.
- **Qxx.** What is the RCM method?
- **Qxx.** What year was RCM published?

A new way to solve the system

So far we have solved the system exactly. But there are situations were a "good enough" solution can be obtained with less work than obtaining the full/complete solution. Which method to choose depends on the matrix A.

- **Q29.** How many arithmetic operations are needed to solve a linear system by forward and backward substitution, once the triangular factors L and U are computed?

- **Q31.** How many arithmetic operations are needed to multiply a $m \times n$ matrix by a $n \times p$ matrix? Compare with the count for Gaussian elimination!

- **Qxx.** What is the Cholesky method? What is it used for?

- We may transform the matrix into a graph. We can then use graph theory for optiizing:

- **Q38.** Describe how a graph defines a matrix and vice versa.

- **Q39.** How does fill-in occur when one does Gaussian elimination on a sparse matrix. Describe it in both matrix and graph terms.

- We then have to agree on a way to measure the error and/or "optimize" the solution.
- We measure the size of an array using a norm.
- Optimzing in euclidian norm will give another solution than optimzing in max-norm or inf-norm.

Quiz

- Given the three data points $x = (1, 2, 3)$ and $y = (2, 3, 5)$. Determine the optimal line if the norm is chosen as

◦ $||r||_2$

◦ $||r||_\infty$

◦ $||r||_1$

- **Qxx.** What is meant by Equivalent norms?

Also, when we get an approximate solution, we get a non-zero residual. First we see that a small residual not neccessarily means a small error.

When trying to solve

$$\begin{pmatrix} 0.324 & 0.973 \\ 0.273 & 0.819 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.649 \\ 0.546 \end{pmatrix}$$

two persons got the different solutions

$$\hat{x} = \begin{pmatrix} -11.128 \\ 4.374 \end{pmatrix} \quad \text{and} \quad \tilde{x} = \begin{pmatrix} -0.936 \\ 0.876 \end{pmatrix}$$

Which solution should we use?

The residuals in the two cases are

$$\hat{r} = A\hat{x} - b = \begin{pmatrix} 0.0014 \\ -0.0016 \end{pmatrix} \quad \text{and} \quad \tilde{r} = A\tilde{x} - b = \begin{pmatrix} -0.0999 \\ -0.0841 \end{pmatrix}$$

with

$$||\hat{r}||_2 = 0.0022 \quad \text{and} \quad ||\tilde{r}||_2 = 0.1306$$

$$\hat{x} = \begin{pmatrix} -11.128 \\ 4.374 \end{pmatrix} \quad \text{and} \quad \tilde{x} = \begin{pmatrix} -0.936 \\ 0.876 \end{pmatrix}$$

$$||\hat{r}||_2 = 0.0022 \quad \text{and} \quad ||\tilde{r}||_2 = 0.1306$$

In fact, the true solution to the system is $x = (\, -1 \quad 1\,)^T$ giving

$$\hat{e} = \hat{x} - x = \begin{pmatrix} -10.128 \\ 3.374 \end{pmatrix} \quad \text{and} \quad \tilde{e} = \tilde{x} - x = \begin{pmatrix} 0.064 \\ -0.124 \end{pmatrix}$$

with

$$||\hat{e}||_2 = 10.675 \quad \text{and} \quad ||\tilde{e}||_2 = 0.1395$$

Remember that

$$Ay = p, Ax = b \rightarrow A * (y - x) = (p - b) \leftrightarrow Ae = r$$

$$||r|| \leq ||A|| \cdot ||e||$$

$$||e|| \leq ||A^{-1}|| \cdot ||r||$$

$$||A||_2 = 1.3405 \qquad ||A^{-1}||_2 = 4910.3$$

- We must then realise that Minimizing the residual might/will give another solution that minizing the error.
- **Qxx.** Is this acceptable? Could we by a solution that has a large error but a small residual?

Least Squares Method

- by Normal Eqs (hi cond),
- QR-decomposition (2x work), or
- SVD (more work, chose rank!).
- Morse Pseudo inverse, choose dimension depending on given/wanted accuracy. (avoiding the very ill-conditioned, nearly not-linearly-independant problems)
- Iterative methods:

- Static methods (Jacobi, Gauss-Seidel)
- Krylov space methods - Conjugate Gradient Method
- Preconditioning
- 3 programming paradigms

EIGENVALUES:

- Methods are divided into two cathegories: direct & iterative
- Similarity transformations, into simpler form
- **Q56.** What is meant by two matrices being similar?
- **Q57.** Show that two similar matrices have the same set of eigenvalues. How are the eigenvectors related?
- Schur theorem

- Power method, inverse power method
- Householder algoritm
- Arnoldi
- Lanczos
- **Qxx.** Why can Lanzcos alg manage much bigger matrices than Arnoldi method?