Lecture 5, Sci. Comp. for DPhil Students

Nick Trefethen, Tuesday 29.10.19

Today

II. Dense linear algebra

- II.1 Matrices, vectors and expansions
- II.2 Orthogonal vectors and matrices

Handout

- `m07_expansions.m` and `m08_leastsquares.m`
- Assignment 2

Announcements

Assignment 2 due is next Tuesday.

These lectures go by easily and I hope are entertaining enough. Nevertheless we are covering important ground fast. I'd like to emphasize that if you are serious about learning this material, you must take time to study lecture notes, M-files, and reading materials:

`https://courses.maths.ox.ac.uk/node/45032`

Four excellent references on iterative linear systems and eigenvalues, available online: Saad 1 & 2, Templates 1 & 2. See the Books and Journals handout or `tools.html` for links.

II. Dense linear algebra

We've had four lectures on large-scale sparse linear algebra: matrices of dimensions in the tens of thousands or more.

For the next four lectures, I want to scale back to the fundamental notions and algorithms of dense numerical linear algebra. You may think you know this material already, but I want to emphasise certain points of view that may not be so familiar to you.

The ways of thinking that we will talk about have proved fruitful for linear and nonlinear problems all across science and engineering.

Good reading: Trefethen & Bau chapters 1–3, pp. 3–24

II.1 Matrices, vectors and expansions

I want you to think about matrix-vector products columnwise:

```
|  |  |   |  |     | x |       |   |
|  |  |   |  |     |  1|       |   |
|  |  |   |  |     |   |       |   |
|  |  |   |  |     | : |       |   |
|  |  |   |  |     |   |       |   |
|  |  |   |  |     | x |       |   |
|a |a |...|a |     |  n|   =   | b |
| 1| 2|   | n|                  |   |
|  |  |   |  |                  |   |
|  |  |   |  |                  |   |
|  |  |   |  |                  |   |
|  |  |   |  |                  |   |
|  |  |   |  |                  |   |


      A              x      =      b
```

$a_j = j$th column of A, $x_j = j$th entry of x

b is a linear combination of the columns of A with coefficients $\{x_j\}$:

$$b = \sum_{j=1}^{n} x_j a_j$$

The **range** or **column space** of A is the set of all such linear combinations.

range(A) is a vector space of dimension $\leq n$.

Its dimension is called the **rank** or **column rank** of A.

rank(A) $= n$: A has **full rank**

rank(A) $< n$: A is **rank-deficient**

We may ask: given b, is there a coefficient vector x such that Ax is equal to b, or close to b?

Suppose A is square and nonsingular, i.e., of full rank n.

Then $Ax = b$ has a unique solution for each b: $x = A^{-1}b$.

$(b_1, \ldots, b_n)^T$: data vector

$(x_1, \ldots, x_n)^T$: coefficient vector

Thus if b is a data vector, $A^{-1}b$ is a coefficient vector.

You should think of this every time you see an expression $A^{-1}b$.

Solution of system of equations $\Longleftrightarrow$ *coefficient vector for an expansion*

On the other hand suppose A has more rows than columns.

Then unless b happens to be in range(A), $Ax = b$ has no solution.

However, one can still look for approximations $Ax \approx b$.

This will lead us to **least-squares**.

We can go further. Suppose each column of A is a function of a continuous variable t, and likewise b. We say that A is an $\infty \times n$ **quasimatrix**.

Thus $a_j = a_j(t)$ and $b = b(t)$.

range(A) is still defined (a vector space of functions of t).

rank(A) is still defined ($\leq n$).

Again $Ax = b$ probably doesn't have a solution, but again it makes sense to look for approximations.

The Chebfun software system starts from these observations (see `www.chebfun.org`). We'll talk a bit about Chebfun later.

`[m07_expansions.m]`

II.2 Orthogonal vectors and matrices

The **inner product** of two vectors v and w is $v^T w = \sum v_j w_j$.

(For functions $v(t)$ and $w(t)$ this generalizes to an integral.)

v and w are **orthogonal** if $v^T w = 0$.

The **norm** or **2-norm** of v is $\|v\| = \|v\|_2 = (v^T v)^{1/2}$.

A set of vectors $\{q_j\}$ are **orthogonal** if they are pairwise orthogonal.
They are **orthonormal (ON)** if $q_j^T q_k = 1$ for $j = k$ and 0 for $j \neq k$.

If $\{q_1, \ldots, q_n\}$ are ON, we get expansion coefficients via inner products:

$$b = \sum_j x_j q_j \iff x_j = q_j^T b$$

or, written via matrices,

```
|   |   |     |   |      | x |         |    |
|   |   |     |   |      |  1|         |    |
|   |   |     |   |      |   |         |    |
|   |   |     |   |      | : |         |    |
|   |   |     |   |      |   |         |    |
|q  |q  |     |q  |      | x |         | b  |
|  1|  2|...  |  n|      |  n|    =    |    |          ( Q x = b )
|   |   |     |   |                    |    |
|   |   |     |   |                    |    |
|   |   |     |   |                    |    |
|   |   |     |   |                    |    |
|   |   |     |   |                    |    |
```

```
| x |        |              T              |  |   |
|  1|        |             q               |  |   |
|   |        |              1              |  |   |
|   |   =    |  - - - - - - - - - - -      |  |   |
| : |        |             :               |  |   |
|   |        |  - - - - - - - - - - -      |  |   |
|   |        |              T              |  | b |
| x |        |             q               |  |   |
|  n|        |              n              |  |   |
                                              |   |
                                              |   |
                    T                         |   |
          ( x = Q b )                         |   |
```

Note that we have used here a rectangular matrix Q with ON columns. There is no standard term for such matrices. We've used the equation

$$Q^T Q = I \qquad (n \times m) \times (m \times n) = (n \times n)$$

An **orthogonal** matrix is a *square* matrix with ON columns.

Suppose Q is orthogonal. Then

```
|            T            |   |   |   |       |   |
|           q             |   |   |   |       |   |
|            1            |   |   |   |       |   |
| - - - - - - - - -       |   |   |   |       |   |
|                         |   |   |   |       |   |
|           :             |   | q | q |       | q |    =   I
|           :             |   |  1|  2| . . . |  n|
|                         |   |   |   |       |   |
| - - - - - - - - -       |   |   |   |       |   |
|            T            |   |   |   |       |   |
|           q             |   |   |   |       |   |
|            n            |   |   |   |       |   |
```

That is, $Q^T Q = I$, or $Q^T = Q^{-1}$.

For such a matrix, we get expansion coefficients via $x = Q^T b = Q^{-1} b$:

```
|   |        |            T            |   |   |
|   |        |           q             |   |   |
|   |        |            1            |   |   |
|   |        | - - - - - - - -         |   |   |
|   |        |                         |   |   |
|   |        |           :             |   |   |
| x |   =    |           :             |   | b |
|   |        |                         |   |   |
|   |        | - - - - - - - -         |   |   |
|   |        |            T            |   |   |
|   |        |           q             |   |   |
|   |        |            n            |   |   |
```

Entrywise, this is obvious: $x_j = q_j^T b$.

Proposition. *If Q is orthogonal, then $\|Qx\| = \|x\|$.*

Proof.

$$\|Qx\|^2 = (Qx)^T(Qx) = x^T Q^T Q x = x^T x = \|x\|^2$$

Suppose A is $m \times n$ with $m > n$.

A **least-squares solution** to $Ax = b$ is a vector x such that

$$\|Ax - b\| = \text{minimum} \qquad (*)$$

Let's draw a picture (which is supposed to show a right angle):

```
          b
          *
              .             /
                  .        /
                      .   /
                         * Ax
                        /
                       /
                      /
                     /                  m
                    /                  R
-------------------/----------------------
                  /
                 /  0
                /
               /
              /  range(A)
             /
```

It's clear geometrically, and easily proved, that $(*)$ is equivalent to the condition that the **residual** $b - Ax$ is orthogonal to range(A).

That is, for each column a_j of A, $a^T(b - Ax) = 0$. Equivalently,

$$A^T(b - Ax) = 0.$$

These are known as the **normal equations,**

$$A^T b = A^T A x.$$

In Matlab, `x = A\b` computes a least-squares solution if $m > n$. We'll see how it does it in the next lecture.

`[m08_leastsquares.m]`

(If extra time, Chebfun quasimatrices)