Gram-Schmidt process and QR factorization 3/6/2001 Math 21b, O. Knill

Homework for Friday, March 9, 2001: Section 4.2, 2,14,16,34,40,42

Motivation. The Gram-Schmidt process is an algorithm to build from an arbitrary basis an **orthonormal basis**. Why do we care to have an orthonormal basis?

- The process of producing an orthonormal basis from a matrix A with column vectors v_j will be associated to a factorization $A = QR$, which helps to solve linear equations. In physical problems like in astrophysics, the numerical methods to simulate the problems one needs to invert huge matrices in every time step of the evolution. (The reason why this is necessary sometimes is to assure the numerical method is stable (implicit methods)).
- An orthonormal basis looks like the standard basis $v_1 = (1, 0, ..., 0), \dots, v_n = (0, 0, ..., 1)$. We actually will see that one can turn an orthonormal basis into a standard basis or a mirror of the standard basis.
- For many physical problems like in quantum mechanics or dynamics, matrices are **symmetric** $A^* = A$, where $A^*_{ij} = A_{ji}$. For such matrices, there is a natural orthonormal basis which has the nice property that $A\phi_i = \lambda_i \phi_i$. The ϕ_i will be called eigenvectors and the λ_i are the eigenvalues (they appear later in the course).
- The **formula for the projection** onto a linear subspace V simplifies if we have an orthonormal basis w_j in V. We have then
$$\mathrm{proj}_V(x) = (w_1 \cdot x)w_1 + \cdots + (w_n \cdot x)w_n \ .$$
- Just for practical reasons, having an orthonormal basis simplifies life - partly because the presence of many zeros $w_j \cdot w_i = 0$ makes computations easier. This is especially the case for problem with some symmetry.
- There is more behind the QR factorization: if we form $A = QR$ in then $A_1 = RQ$, the new matrix $A_1 = Q^{-1}AQ$ shares many properties of A (like the eigenvalues about which we will learn in a few weeks). When iterating this procedure $A \to A_1$ one is lead to interesting topics in differential equations (Toda lattice).

The Gram-Schmidt process.
Leb $v_1, ..., v_n$ be a basis. Let $u_1 = v_1$ and $w_1 = u_1/||u_1||$. The Gram-Schmidt process recursively constructs from the already constructed orthonormal set $w_1, ..., w_{i-1}$ which spans a linear space V_{i-1} the new vector $u_i = (v_i - \mathrm{proj}_{V_{i-1}}(v_i))$ which is orthogonal to V_{i-1}, and then normalizing u_i to to get $w_i = u_i/||u_i||$. The vectors u_i are orthogonal to the linear space V_{i-1}.

A simple example.

Find an orthonormal basis for $v_1 = \left|\begin{array}{c} 2 \\ 0 \\ 0 \end{array}\right|$, $v_2 = \left|\begin{array}{c} 1 \\ 3 \\ 0 \end{array}\right|$ and $v_3 = \left|\begin{array}{c} 1 \\ 2 \\ 5 \end{array}\right|$.

Solution.

1. $w_1 = v_1/||v_1|| = \left|\begin{array}{c} 1 \\ 0 \\ 0 \end{array}\right|$.

2. $u_2 = (v_2 - \mathrm{proj}_{V_1}(v_2) = v_2 - (w_1 \cdot v_2)w_1, = \left|\begin{array}{c} 0 \\ 3 \\ 0 \end{array}\right|$. $w_2 = u_2/||u_2|| = \left|\begin{array}{c} 0 \\ 1 \\ 0 \end{array}\right|$.

3. $u_3 = (v_3 - \mathrm{proj}_{V_2}(v_3)) = v_3 - (w_1 \cdot v_3)w_1 - (w_2 \cdot v_3)w_2 = \left|\begin{array}{c} 0 \\ 0 \\ 5 \end{array}\right|$, $w_3 = u_3/||u_3|| = \left|\begin{array}{c} 0 \\ 0 \\ 1 \end{array}\right|$.

The Factorization.
The formulas can be written as

$$v_1 = ||v_1||w_1 = r_{11}w_1$$

$$\cdots$$

$$v_i = (w_1 \cdot v_i)w_1 + \cdots + (w_{i-1} \cdot v_i)w_{i-1} + ||u_i||w_i = r_{i1}w_1 + \cdots + r_{ii}w_i$$

$$\cdots$$

$$v_n = (w_1 \cdot v_n)w_1 + \cdots + (w_{n-1} \cdot v_n)w_{n-1} + ||u_n||w_n = r_{n1}w_1 + \cdots + r_{nn}w_n$$

which means in matrix form

$$A = \begin{vmatrix} | & | & \cdot & | \\ v_1 & \cdots & \cdot & v_m \\ | & | & \cdot & | \end{vmatrix} = \begin{vmatrix} | & | & \cdot & | \\ w_1 & \cdots & \cdot & w_m \\ | & | & \cdot & | \end{vmatrix} \begin{vmatrix} r_{11} & r_{12} & \cdot & r_{1m} \\ 0 & r_{22} & \cdot & r_{2m} \\ 0 & 0 & \cdot & r_{mm} \end{vmatrix} = QR$$

where A and Q are $n \times m$ matrices and R is a $m \times m$ matrix.

> Any matrix A can be decomposed as $A = QR$, where Q has as columns orthonormal vectors and R is an upper triangular square matrix.

Back to the example.

The matrix with the vectors v_1, v_2, v_3 is $A = \begin{vmatrix} 2 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 5 \end{vmatrix}$.

$v_1 = ||v_1||w_1$
$v_2 = (w_1 \cdot v_2)w_1 + ||u_2||w_2$
$v_3 = (w_1 \cdot v_3)w_1 + (w_2 \cdot v_3)w_2 + ||u_3||w_3$,
so that $Q = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$ and $R = \begin{vmatrix} 2 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 5 \end{vmatrix}$. Since the matrix A was trigonal, the QR decomposition is trivial.

How to memorize this?.
While building the matrix R we keep track of the vectors u_i during the Gram-Schmidt procedure. At the end you have vectors u_i, v_i, w_i and the matrix R has the $||u_i||$ in the diagonal as well as the dot products $w_i \cdot v_j$ in the upper right triangle.

History of the Gram-Schmidt process.
The recursive formulae of the process were stated by Erhard Schmidt (1876-1959) in (Math. Annalen 63 (1907) 433-476). Implicitly the essence of the formulae were in a 1883 paper of J.P.Gram (Crelle 94 (1883) 41-73), which Schmidt mentions in a footnote. Also Gram seems not have been the first to note it. The process seems to be a result of Laplace (1749-1827) and it was essentially used by Cauchy (1789-1857) in 1836 already.

The pictures below show from the left to the right, Gram, Schmidt, Laplace and Cauchy.