Gram-Schmidt process and QR factorization 3/8/2002 — Math 21b, O. Knill

Homework: Section 5.2, 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 $\vec{v}_j$ 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 $\vec{v}_1 = (1, 0, ..., 0), \dots, \vec{v}_n = (0, 0, ..., 1)$. We actually will see that one can turn an orthonormal basis into a standard basis or a mirror of the standard basis.
- If A is written as $A = QR$, then the inverse of A can be found easier, because the inverse of Q and the inverse of R are easy to obtain. See later.
- 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 $A\vec{v}_i = \lambda_i \vec{v}_i$. The $\vec{v}_i$ will be called eigenvectors and the λ_i are the eigenvalues. These objects will appear later in the course.
- The **formula for the projection** onto a linear subspace V simplifies with an orthonormal basis $\vec{v}_j$ in V:

$$\mathrm{proj}_V(\vec{x}) = (\vec{v}_1 \cdot \vec{x})\vec{w}_1 + \cdots + (\vec{w}_n \cdot \vec{x})\vec{w}_n \,.$$

- For practical reasons, having an orthonormal basis simplifies life - partly because the presence of many zeros $\vec{w}_j \cdot \vec{w}_i = 0$ makes computations easier. This is especially the case for problems with 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 (an example is the Toda lattice).

GRAM-SCHMIDT PROCESS.
Let $\vec{v}_1, ..., \vec{v}_n$ be a basis. Let $\vec{u}_1 = \vec{v}_1$ and $\vec{w}_1 = \vec{u}_1/||\vec{u}_1||$. The Gram-Schmidt process recursively constructs from the already constructed orthonormal set $\vec{w}_1, ..., \vec{w}_{i-1}$ which spans a linear space V_{i-1} the new vector $\vec{u}_i = (\vec{v}_i - \mathrm{proj}_{V_{i-1}}(\vec{v}_i))$ which is orthogonal to V_{i-1}, and then normalizing $\vec{u}_i$ to to get $\vec{w}_i = \vec{u}_i/||\vec{u}_i||$. The vectors $\vec{u}_i$ are orthogonal to the linear space V_{i-1}.

EXAMPLE.
Find an orthonormal basis for $\vec{v}_1 = \begin{vmatrix} 2 \\ 0 \\ 0 \end{vmatrix}$, $\vec{v}_2 = \begin{vmatrix} 1 \\ 3 \\ 0 \end{vmatrix}$ and $\vec{v}_3 = \begin{vmatrix} 1 \\ 2 \\ 5 \end{vmatrix}$.

SOLUTION.

1. $\vec{w}_1 = \vec{v}_1/||\vec{v}_1|| = \begin{vmatrix} 1 \\ 0 \\ 0 \end{vmatrix}$.

2. $\vec{u}_2 = (\vec{v}_2 - \mathrm{proj}_{V_1}(\vec{v}_2)) = \vec{v}_2 - (\vec{w}_1 \cdot \vec{v}_2)\vec{w}_1, = \begin{vmatrix} 0 \\ 3 \\ 0 \end{vmatrix}$. $\vec{w}_2 = \vec{u}_2/||\vec{u}_2|| = \begin{vmatrix} 0 \\ 1 \\ 0 \end{vmatrix}$.

3. $\vec{u}_3 = (\vec{v}_3 - \mathrm{proj}_{V_2}(\vec{v}_3)) = \vec{v}_3 - (\vec{w}_1 \cdot \vec{v}_3)\vec{w}_1 - (\vec{w}_2 \cdot \vec{v}_3)\vec{w}_2 = \begin{vmatrix} 0 \\ 0 \\ 5 \end{vmatrix}$, $\vec{w}_3 = \vec{u}_3/||\vec{u}_3|| = \begin{vmatrix} 0 \\ 0 \\ 1 \end{vmatrix}$.

2. EXAMPLE.
$\vec{v}_1 = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$ $\vec{v}_2 = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}$ $\vec{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$ leads to orthonormal basis $\vec{w}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ $\vec{w}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$ $\vec{w}_3 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}$

QR FACTORIZATION.
The formulas can be written as

$$\vec{v}_1 = ||\vec{v}_1||\vec{w}_1 = r_{11}\vec{w}_1$$

$$\cdots$$

$$\vec{v}_i = (\vec{w}_1 \cdot \vec{v}_i)\vec{w}_1 + \cdots + (\vec{w}_{i-1} \cdot \vec{v}_i)\vec{w}_{i-1} + ||\vec{u}_i||\vec{w}_i = r_{i1}\vec{w}_1 + \cdots + r_{ii}\vec{w}_i$$

$$\cdots$$

$$\vec{v}_n = (\vec{w}_1 \cdot \vec{v}_n)\vec{w}_1 + \cdots + (\vec{w}_{n-1} \cdot \vec{v}_n)\vec{w}_{n-1} + ||\vec{u}_n||\vec{w}_n = r_{n1}\vec{w}_1 + \cdots + r_{nn}\vec{w}_n$$

which means in matrix form

$$A = \left|\begin{array}{cccc} | & | & \cdot & | \\ \vec{v}_1 & \cdots & \cdot & \vec{v}_m \\ | & | & \cdot & | \end{array}\right| = \left|\begin{array}{cccc} | & | & \cdot & | \\ \vec{w}_1 & \cdots & \cdot & \vec{w}_m \\ | & | & \cdot & | \end{array}\right| \left|\begin{array}{cccc} r_{11} & r_{12} & \cdot & r_{1m} \\ 0 & r_{22} & \cdot & r_{2m} \\ 0 & 0 & \cdot & r_{mm} \end{array}\right| = QR$$

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.

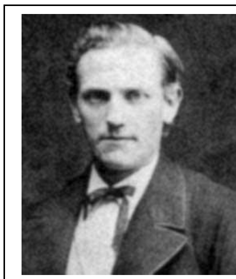
1. EXAMPLE. The matrix with the vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3$ is $A = \left|\begin{array}{ccc} 2 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 5 \end{array}\right|$.

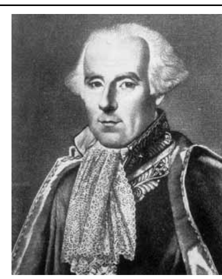
$\vec{v}_1 = ||\vec{v}_1||\vec{w}_1$
$\vec{v}_2 = (\vec{w}_1 \cdot \vec{v}_2)\vec{w}_1 + ||\vec{u}_2||\vec{w}_2$
$\vec{v}_3 = (\vec{w}_1 \cdot \vec{v}_3)\vec{w}_1 + (\vec{w}_2 \cdot \vec{v}_3)\vec{w}_2 + ||\vec{u}_3||\vec{w}_3$,
so that $Q = \left|\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right|$ and $R = \left|\begin{array}{ccc} 2 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 5 \end{array}\right|$.

PRO MEMORIA.
The matrix R has the $||\vec{u}_i||$ in the diagonal and the dot products $[R]_{ij} = \vec{w}_i \cdot \vec{v}_j$ in the upper right triangle.

3. EXAMPLE. Make the QR decomposition of $A = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$. $\vec{w}_1 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$. $\vec{u}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$.
$\vec{w}_2 = \vec{u}_2$. $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = QR$.

SOME HISTORY. The recursive formulae of the process were stated by Erhard Schmidt (1876-1959) in 1907 Implicitly the essence of the formulae were in a 1883 paper of J.P.Gram in 1883 which Schmidt mentions in a footnote. The process seems to be a result of Laplace (1749-1827). It was already also used by Cauchy (1789-1857) in 1836.






Gram Schmidt Laplace Cauchy