GRAM SCHMIDT AND QR FACTORIZATION   Math 21b, O. Knill

GRAM-SCHMIDT PROCESS.
Let $\vec{v}_1, ..., \vec{v}_n$ be a basis in V. 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 $\vec{w}_i = \vec{u}_i/||\vec{u}_i||$. Each vector $\vec{w}_i$ is orthonormal to the linear space V_{i-1}. The vectors $\{\vec{w}_1, .., \vec{w}_n\}$ form an orthonormal basis in V.
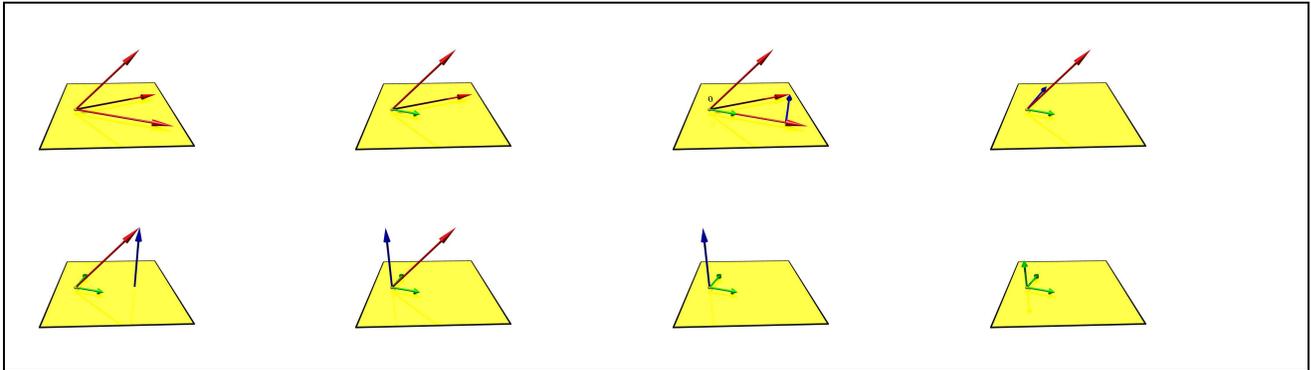
EXAMPLE.
Find an orthonormal basis for $\vec{v}_1 = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$ and $\vec{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$.

SOLUTION.

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

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

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

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

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

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

> THE GRAM-SCHMIDT PROCESS PROVES: Any matrix A with linearly independent columns $\vec{v}_i$ can be decomposed as $A = QR$, where Q has orthonormal column vectors and where R is an upper triangular square matrix. The matrix Q has the orthonormal vectors $\vec{w}_i$ in the columns.

BACK TO THE EXAMPLE. The matrix with the vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3$ is $A = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 5 \end{bmatrix}$.

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

PRO MEMORIA.
While building the matrix R we keep track of the vectors u_i during the Gram-Schmidt procedure. At the end you have vectors $\vec{u}_i, \vec{v}_i, \vec{w}_i$ and the matrix R has $\boxed{||\vec{u}_i||}$ in the diagonal as well as the dot products $\boxed{\vec{w}_i \cdot \vec{v}_j}$ in the upper right triangle where $i < j$.

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

WHY do we care to have an orthonormal basis?

- An orthonormal basis looks like the standard basis $\vec{v}_1 = (1, 0, ..., 0), ..., \vec{v}_n = (0, 0, ..., 1)$. Actually, we will see that an orthonormal basis into a standard basis or a mirror of the standard basis.

- The Gram-Schmidt process is tied to the factorization $A = QR$. The later 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. Inverting $A^{-1} = R^{-1}Q^{-1}$ is easy because R and Q are easy to invert.

- For many physical problems like in quantum mechanics or dynamical systems, matrices are **symmetric** $A^* = A$, where $A_{ij}^* = \overline{A}_{ji}$. For such matrices, there will a natural orthonormal basis.

- The **formula for the projection** onto a linear subspace V simplifies with an orthonormal basis $\vec{v}_j$ in V: $\text{proj}_V(\vec{x}) = (\vec{v}_1 \cdot \vec{x})\vec{w}_1 + \cdots + (\vec{w}_n \cdot \vec{x})\vec{w}_n$.

- An orthonormal basis simplifies compuations due to the presence of many zeros $\vec{w}_j \cdot \vec{w}_i = 0$. This is especially the case for problems with symmetry.

- The Gram Schmidt process can be used to define and construct classes of classical polynomials, which are important in physics. Examples are Chebyshev polynomials, Laguerre polynomials or Hermite polynomials.

- QR factorization allows fast computation of the determinant, least square solutions $R^{-1}Q^{-1}\vec{b}$ of overdetermined systems $A\vec{x} = \vec{b}$ or finding eigenvalues - all topics which will appear later.

SOME HISTORY.

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

       Gram   Schmidt   Laplace   Cauchy