QR-factorization

The purpose of this note is to explain QR-factorization. The point of departure is SIF, the chapter on Orthogonality, the section on Orthogonal vecotrs (6.2 eller 7.2), *Let A be an $m \times n$ matrix such that the column vectors of A are linearly independent. There is an orthogonal matrix Q and an upper triangular matrix R such that $A = QR$.*

The argument for this, and the algorithm to find Q and R is via the Gram Schmidt algorithm and goes as follows: Let the columns of A be $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$.

The Gram Schmidt procedure, which finds an orthogonal basis $\mathbf{v}_1, \dots \mathbf{v}_n$ with the same span as $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$, is:

- $\mathbf{v}_1 = \mathbf{a}_1$,
- $\mathbf{v}_2 = \mathbf{a}_2 - \frac{\mathbf{a}_2 \cdot \mathbf{v}_1}{|\mathbf{v}_1|^2}\mathbf{v}_1$ (the projection of $\mathbf{a}_2$ on $\mathbf{v}_1$ is subtracted from $\mathbf{a}_2$),
- $\mathbf{v}_3 = \mathbf{a}_3 - (\frac{\mathbf{a}_3 \cdot \mathbf{v}_2}{|\mathbf{v}_2|^2}\mathbf{v}_2 + \frac{\mathbf{a}_3 \cdot \mathbf{v}_1}{|\mathbf{v}_1|^2}\mathbf{v}_1)$ (the projection of $\mathbf{a}_3$ on span$(\mathbf{v}_1, \mathbf{v}_2)$ is subtracted from $\mathbf{a}_3$).
- In general $\mathbf{v}_j = \mathbf{a}_j - (\frac{\mathbf{a}_j \cdot \mathbf{v}_{j-1}}{|\mathbf{v}_{j-1}|^2}\mathbf{v}_{j-1} + \frac{\mathbf{a}_j \cdot \mathbf{v}_{j-2}}{|\mathbf{v}_{j-2}|^2}\mathbf{v}_{j-2} + \cdots + \frac{\mathbf{a}_j \cdot \mathbf{v}_1}{|\mathbf{v}_1|^2}\mathbf{v}_1)$

The columns of the matrix Q are $\mathbf{q}_j = \frac{1}{|\mathbf{v}_j|}\mathbf{v}_j$. They are orthogonal, since this is what comes out of the Gram-Schmidt proces.

From the general expression for $\mathbf{v}_j$ we get

$$\mathbf{a}_j = \mathbf{v}_j + (\frac{\mathbf{a}_j \cdot \mathbf{v}_{j-1}}{|\mathbf{v}_{j-1}|^2}\mathbf{v}_{j-1} + \frac{\mathbf{a}_j \cdot \mathbf{v}_{j-2}}{|\mathbf{v}_{j-2}|^2}\mathbf{v}_{j-2} + \cdots + \frac{\mathbf{a}_j \cdot \mathbf{v}_1}{|\mathbf{v}_1|^2}\mathbf{v}_1)$$

In particular, $\mathbf{a}_j$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_j$ and hence of $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_j$.

$$\mathbf{a}_j = |\mathbf{v}_j|\mathbf{q}_j + (\frac{\mathbf{a}_j \cdot \mathbf{v}_{j-1}}{|\mathbf{v}_{j-1}|}\mathbf{q}_{j-1} + \frac{\mathbf{a}_j \cdot \mathbf{v}_{j-2}}{|\mathbf{v}_{j-2}|}\mathbf{q}_{j-2} + \cdots + \frac{\mathbf{a}_j \cdot \mathbf{v}_1}{|\mathbf{v}_1|}\mathbf{q}_1)$$

It follows that $A = QR$, where R is upper triangular and $r_{ij} = \frac{\mathbf{a}_j \cdot \mathbf{v}_i}{|\mathbf{v}_i|}$. In particular, the diagonal entrances are $r_{ii} = |\mathbf{v}_i|$, since $\mathbf{a}_i \cdot \mathbf{v}_i = \mathbf{v}_i \cdot \mathbf{v}_i$ because of orthogonality.

FREDRIK BAJERSVEJ 7G
9220 AALBORG ØST
9940 8848
-
FAJSTRUP@MATH.AAU.DK
PEOPLE.MATH.AAU.DK/~FAJSTRUP