6 QR decomposition

Monday, September 21, 2020 11:15 PM

Def 11.8 Given any $A \in \mathbb{R}^{n \times n}$, a *QR-decomposition* of A is any pair of $n \times n$ matrices (Q, R), where Q is an orthogonal matrix and R is an upper triangular matrix s.t. $A = QR$.

(Note that if A is not invertible, then some diagonal entry in R must be 0.)

Prop. 11.14/11.16 Given any invertible $A \in \mathbb{R}^{n \times n}$, then $\exists$ an orthogonal Q and upper triangular R with pos. diag. entries s.t. $A = QR$.

proof. Let $A = (A^1 \cdots A^n)$, so $A^i \in \mathbb{R}^n$ are cols of A.

If A is invertible, then A^i's are lin. ind., so we can produce an orthonormal basis using Gram-Schmidt orthonormalization.

let $Q'^1 = A^1$, $Q^1 = \dfrac{Q'^1}{\|Q'^1\|}$.

and $Q'^{k+1} = A^{k+1} - \sum_{i=1}^{k} (A^{k+1} \cdot Q^i) Q^i$, $Q^{k+1} = \dfrac{Q'^{k+1}}{\|Q'^{k+1}\|}$.

We can rewrite those equations:

$$A^1 = \|Q'^1\| Q^1$$

$$\vdots$$

$$A^j = \sum_{k=1}^{j-1} (A^j \cdot Q^k) Q^k + \|Q'^j\| Q^j$$

$$\vdots$$

$$A^n = \sum_{k=1}^{n} (A^n \cdot Q^k) Q^k + \|Q'^n\| Q^n$$

Let $r_{kk} = \|Q'^k\|$ and $r_{ij} = A^j \cdot Q^i$, $1 \le k \le n$, $2 \le j \le n$, $1 \le i \le j-1$.

Then $A^j = \sum_{k=1}^{j-1} r_{kj} Q^k + r_{jj} Q^j$

So $a_{ij} = \sum_{k=1}^{j-1} r_{kj} q_{ik} + r_{jj} q_{ij}$ for $2 \le j \le n$, $1 \le i \le j-1$.

Note:

$$\begin{bmatrix} q_{11} & \cdots & q_{1n} \\ \vdots & \ddots & \vdots \\ q_{n1} & \cdots & q_{nn} \end{bmatrix} \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ & \ddots & \vdots \\ 0 & & r_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} q_{11} r_{11} & q_{11} r_{12} + q_{12} r_{22} & \cdots & \sum_{k=1}^{n-1} q_{1k} r_{kn} + q_{1n} r_{nn} \\ \vdots & & & \vdots \end{bmatrix} = (a_{ij})_{i,j}$$

Thus, if we let $Q = (Q^1 \cdots Q^n)$

$R = (r_{ij})$ as given above (and 0 below the diagonal),

$A = QR$.

Aside: The QR decomposition is unique when requiring pos. diag. entries

because if $A = Q_1 R_1 = Q_2 R_2$

$\Rightarrow R_1 R_2^{-1} = Q_1^T Q_2$ is orthogonal

$\Rightarrow R_1 R_2^{-1}$ is diagonal (because upper triangular and orthogonal) with ± 1 entries.

$\Rightarrow Q_2 = Q_1 D$ and $R_2 = D R_1$.

We will see later that we do not require invertibility.

Prop 11.5/11.17 (Hadamard) For any real $n \times n$ matrix $A = (a_{ij})$,

$$|\det(A)| \leq \prod_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij}^2 \right)^{\frac{1}{2}} \quad \text{and} \quad |\det(A)| \leq \prod_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ij}^2 \right)^{\frac{1}{2}}.$$

Equality holds iff either A has a zero row in the left inequality

or a zero col in the right inequality

or the columns of A are orthogonal.

proof. If $\det(A) = 0$, inequality is trivial.

Further, if RHS $= 0$, then some column or row is 0.

If $\det(A) \neq 0$ then $A = QR$, $r_{ii} > 0$, Q orthogonal, R upper triangular.

Further, if RHS = 0, then some column of ... is 0.

If $\det(A) \neq 0$, then $A = QR$, $r_{ii} > 0$, Q orthogonal, R upper triangular.

$\det(Q^T Q) = 1 \implies \det(Q) = \pm 1$.

$$\implies |\det(A)| = |\det(Q)|\,|\det(R)| = \prod_{j=1} r_{jj}.$$

Note $\sum_{i=1}^{n} a_{ij}^2 = \|A^j\|_2^2 = \|QR^j\|_2^2 = \|R^j\|_2^2 = \sum_{i=1}^{n} r_{ij}^2 \geq r_{jj}^2$.

$$\implies |\det(A)| = \prod_{j=1}^{n} r_{jj} \leq \prod_{j=1}^{n} \|R^j\|_2 = \prod_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ij}^2\right)^{\frac{1}{2}}.$$

Can apply same proof to A^T to get other ineq.

Finally, if $\det(A) \neq 0$ and equality holds, then

$A = QR$

$Q^T A = R$

$$r_{jj} = \|A^j\|_2$$

$$\implies r_{jj} = \|Q^T A^j\|_2 = \|R^j\|_2$$

$$\implies r_{ij} = 0 \quad \forall\, i \neq j. \implies R \text{ is diagonal}$$

$$\implies A = QR = \begin{pmatrix} Q^1 & \cdots & Q^n \end{pmatrix} \begin{pmatrix} r_{11} & & 0 \\ & \ddots & \\ 0 & & r_{nn} \end{pmatrix} = \begin{pmatrix} r_{11} Q^1 & \cdots & r_{nn} Q^n \end{pmatrix}.$$

$\implies$ rows of A are orthogonal. $\square$

Prop. 11.16/11.18 (Hadamard) If $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ is symmetric pos. semi-definite then $\det(A) \leq \prod_{i=1}^{n} a_{ii}$.

Moreover, if A is pos def., equality holds iff A is diagonal.

proof. Trivial if $\det(A) = 0$.

Otherwise, let $A = B^T B$, B an upper triangular matrix with pos. diagonal entries. (Cholesky)

Then $\det(A) = \det(B)^2$.

By the prev. Prop., $\det(B) \leq \prod_{j=1}^{n} \left(\sum_{i=1}^{n} b_{ij}^2\right)^{\frac{1}{2}}$.

But $a_{jj} = \sum_{i=1}^{n} b_{ij}^2 \implies \det(B) \leq \prod_{j=1}^{n} \sqrt{a_{jj}}$

$$\implies \det(A) = \det(B)^2 \leq \prod_{j=1}^{n} a_{jj}.$$

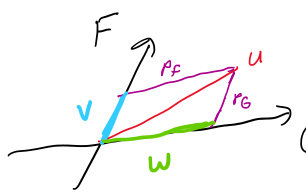
For equality, note that then B has to be orthogonal $\implies B$ is diagonal $\implies A$ is diagonal. $\boxtimes$

So far, we have only shown the QR decomposition for square invertible matrices. Here, we show that you don't need invertibility.

Aside: can even do for rectangular matrices $m \times n$, if $m \geq n$.

Define: Recall that given a vector space E, and subspaces F and G of E, the *direct sum* $E = F \oplus G$ means that $\forall u \in E$, u can be uniquely written as $u = v + w$, where $v \in F$ and $w \in G$.

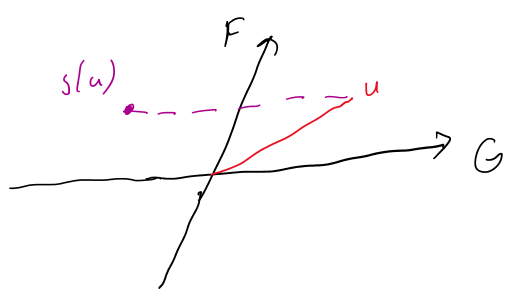
Define: We define the *projections* $p_F : E \to F$ and $p_G : E \to G$ s.t. $p_F(u) = v$ and $p_G(u) = w$, where $u = v + w$.



Note: p_F and p_G are linear, and $p_F^2 = p_F$, $p_G^2 = p_G$, $p_F \circ p_G = p_G \circ p_F = 0$, $p_F + p_G = id$.

Definition 12.1 Given a vector space $E = F \oplus G$, the *symmetry* (or *reflection*) with respect to F and parallel to G is the linear map $s : E \to E$ defined s.t. $s(u) = 2p_F(u) - u$, $\forall u \in E$.

Note: $s(u) = p_F(u) - p_G(u) = u - 2p_G(u)$, $s^2 = id$, $s = id_F$, $s = -id_G$.



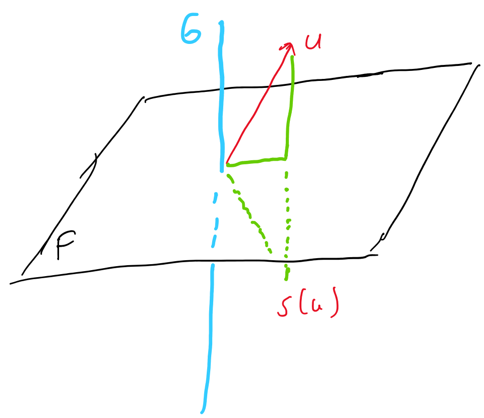
Definition 12.2 Let E be a Euclidean space of finite dim n. For [illegible] orthogonal

Definition 12.2 Let E be a Euclidean space of finite dim n. For any two subspaces F and G, if $E = F \oplus G$ and F and G are orthogonal (i.e. $F = G^{\perp}$), the orthogonal symmetry (or reflection) w.r.t. F and parallel to G is the linear map $s: E \to E$ defined s.t.

$$s(u) = 2p_F(u) - u = p_F(u) - p_G(u), \quad \forall u \in E.$$

When F is a hyperplane, we call s a hyperplane symmetry with respect to F (or reflection about F) and when G is a plane (i.e. $\dim(F) = n-2$), we call s a flip about F.

(recall hyperplanes have dim $n-1$)



Lemma: Let $s: E \to E$ be an orthogonal symmetry w.r.t. F and parallel to G. Then s is an isometry.

proof. $\forall u, v \in E$, $\|u+v\|^2 - \|u-v\|^2 = 4(u \cdot v)$

So if $u \cdot v = 0$, then $\|u+v\| = \|u-v\|$

(intuitively, flipping the sign of a coordinate doesn't change the Euclidean norm)

Recall $u = p_F(u) + p_G(u)$ and $s(u) = p_F(u) - p_G(u)$

Further, F and G are orthogonal, so $p_F(u) \cdot p_G(u) = 0$.

Thus, $\|s(u)\| = \|p_F(u) - p_G(u)\| = \|p_F(u) + p_G(u)\| = \|u\|$. $\blacksquare$

Recall: We can find an orthonormal basis $(e_1, \dots, e_n)$ of E s.t. $(e_1, \dots, e_p)$ is an orthonormal basis of F and $(e_{p+1}, \dots, e_n)$ is an orthonormal basis of G.

Then, the matrix of s in this basis is

$$\begin{pmatrix} I_p & 0 \\ 0 & -I_{n-p} \end{pmatrix}.$$

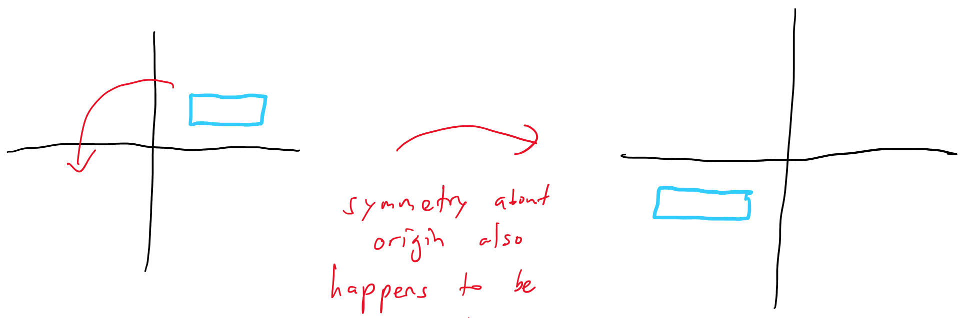
$\Rightarrow \det(s) = (-1)^{n-p}$

Thus, s is a rotation if $n-p$ is even (recall, rotations have $\det = 1$)

If $F = \{0\}$, then $s = -id$, the symmetry w.r.t. the origin

(a rotation if n is even)

(an improper orthogonal transformation if n is odd)



If $F = H$ a hyperplane (i.e. $\dim(F) = n-1$), and s is an orthogonal symmetry w.r.t. F and parallel to G, then $p_G(u) = \lambda w$ for some $\lambda \in G$, $\lambda \neq 0$ orthogonal to H.

Then, because $u \cdot w = \lambda \|w\|^2$

$$\Rightarrow p_G(u) = \frac{(u \cdot w)}{\|w\|^2} w$$

Then $s(u) = u - 2p_G(u) = u - 2\dfrac{(u \cdot w)}{\|w\|^2} w.$

Prop. 12.1: Let E be a finite-dim. Euclidean space and let H be a hyperplane in E. For any nonzero vector w orthogonal to H, the hyperplane reflection s about H is given by

$$s(u) = u - 2\frac{(u \cdot w)}{\|w\|^2} w, \quad u \in E.$$

Of course, we can also write s in matrix form.

Def. 12.3: A Householder matrix is a matrix of the form

$$H = I_n - 2 \cdot \frac{WW^T}{\|W\|^2} = I_n - 2 \cdot \frac{WW^T}{W^T W},$$

where $W \in \mathbb{R}^n$ is a nonzero vector.

Note: H is clearly symmetric.

$$H^T H = \left(I_n - 2 \cdot \frac{WW^T}{\|W\|^2}\right)^2 = I_n - 4 \cdot \frac{WW^T}{\|W\|^2} + 4 \cdot \frac{W(W^T W)W^T}{\|W\|^2} = I_n;$$

so H is orthogonal.

Suppose H is the hyperplane orthogonal to a nonzero vector w, and W is the representation of w in an orthonormal basis $(e_1, \dots, e_n)$.

Then $P_G = \dfrac{WW^T}{W^T W}$ and $P_H = I_n - \dfrac{WW^T}{W^T W}$, so the Householder matrix $H = I_n - 2 \cdot \dfrac{WW^T}{\|W\|^2}$ is exactly the reflection about H under the chosen basis.

$\left(P_G(u) = \dfrac{(u \cdot w)}{\|w\|^2} w \right)$ $\qquad$ $\left(P_G + P_H = \mathrm{Id} \right)$

Aside: Rotations in $\mathbb{R}^3$ can be expressed in this form.

Prop. 12.2 Let E be any nontrivial Euclidean space, $\forall u, v \in E$, if $\|u\| = \|v\|$, then there is a hyperplane H s.t. the reflection s about H maps u to v, and if $u \neq v$, then this reflection is unique.

proof. If $u = v$, then let $u \in H$. Any such H will do.

Otherwise, let $H = \{v - u\}^{\perp}$

$$\Rightarrow s(u) = u - 2\frac{u \cdot (v-u)}{\|v-u\|^2}(v-u) = u + \frac{2\|u\|^2 - 2u \cdot v}{\|v-u\|^2}(v-u)$$

Note $\|v-u\|^2 = \|u\|^2 + \|v\|^2 - 2u \cdot v = 2\|u\|^2 - 2u \cdot v$

$$\Rightarrow s(u) = u + (v-u) = v.$$

Now that we have Householder matrices, we can get another more general proof of QR-decompositions, even for non-invertible matrices.

Prop. 12.3 Let E be a nontrivial Euclidean space of dim n. For any orthonormal basis $(e_1, \dots, e_n)$ and for any n-tuple of vectors $(v_1, \dots, v_n)$, there is a sequence of n isometries $h_1, \dots, h_n$ s.t. h_i is a hyperplane reflection or the identity, and if $(r_1, \dots, r_n)$ are the vectors given by

$$r_j = h_n \circ \cdots h_2 \circ h_1 (v_j),$$

then every r_j is a linear combination of the vectors $(e_1, \dots, e_j)$.
Equivalently, the matrix R whose cols are r_j expressed over $(e_1, \dots, e_n)$ is an upper triangular matrix.
Furthermore, the h_i can be chosen so that the diagonal entries of R are nonnegative.

proof. For $n=1$, $v_1 = \lambda e_1$ for some $\lambda \in \mathbb{R}$.
If $\lambda \geq 0$, let $h_1 = id$.
Else, let $h_1 = -id$, the reflection about the origin.

For $n \geq 2$, we need to find h_1, and then we proceed by induction.

Let $r_{1,1} = \|v_1\|$.
If $v_1 = r_{1,1} e_1$, let $h_1 = id$.
Else, let h_1 be the unique hyperplane reflection taking v_1 to $r_{1,1} e_1$.

$$\left(\begin{array}{l} h_1(v_1) = r_{1,1} e_1 \\ h_1(u) = u - 2 \dfrac{(u \cdot w_1)}{\|w_1\|^2} w_1 , \quad w_1 = r_{1,1} e_1 - v_1 \end{array} \right)$$

Clearly, then $r_1 = \operatorname{span}\{e_1\}$ and $r_{1,1} = \|v_1\| \geq 0$.

Now let's suppose we already have k linear maps $h_1, \dots, h_k$ such that $r_j = h_k \circ \cdots \circ h_1 (v_j)$ is a linear combination of $e_1, \dots, e_j$.

Let $U_k' = \operatorname{span}\{e_1, \dots, e_k\}$

$U_k'' = \operatorname{span}\{e_{k+1}, \dots, e_n\}$, so $E = U_k' \oplus U_k''$.

$U_k'' = \operatorname{span}\{e_{k+1}, \dots, e_n\}$, so $E = U_k' \oplus U_k''$.

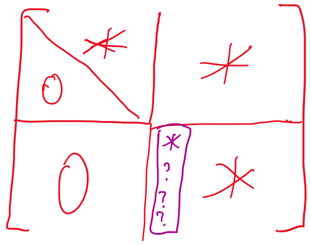
Let $u_{k+1} = h_k \circ \cdots \circ h_1(v_{k+1})$. (the application of the first k reflections to v_{k+1})

Then $u_{k+1} = u_{k+1}' + u_{k+1}''$, where $u_{k+1}' \in U_k'$ and $u_{k+1}'' \in U_k''$.

Let $r_{k+1,k+1} = \| u_{k+1}'' \|$.

If $u_{k+1}'' = r_{k+1,k+1} e_{k+1}$, let $h_{k+1} = id$.

(If the part of u_{k+1} not in the span of the first k basis vectors is exactly in the span of the $(k+1)$th basis vector, then we don't have to do anything)

$H^k \cdots H^1 V =$ 

Otherwise, choose the unique hyperplane reflection s.t.

$$h_{k+1}(u_{k+1}'') = r_{k+1,k+1} e_{k+1}.$$

defined by $h_{k+1}(u) = u - \dfrac{2(u \cdot w_{k+1})}{\|w_{k+1}\|^2} w_{k+1}$, $w_{k+1} = r_{k+1,k+1} e_{k+1} - u_{k+1}''$.

This of course corresponds to reflection about the hyperplane H_{k+1} orthogonal to w_{k+1}.

But $u_{k+1}'', e_{k+1} \in U_k''$, so $w_{k+1} \in U_k''$.

$$\Rightarrow H_{k+1} = \{w_{k+1}\}^{\perp} \supseteq U_k'.$$

$\Rightarrow r_1, \dots, r_k$ and u_{k+1}' are invariant under h_{k+1}.

$\Rightarrow h_{k+1}(u_{k+1}) = h_{k+1}(u_{k+1}') + h_{k+1}(u_{k+1}'') = u_{k+1}' + r_{k+1,k+1} e_{k+1} \in \operatorname{span}\{e_1, \dots, e_{k+1}\}$.

Let $r_{k+1} = h_{k+1}(u_{k+1}) \in \operatorname{span}\{e_1, \dots, e_{k+1}\}$.

$= h_{k+1} \circ h_k \circ \cdots h_2 \circ h_1 (v_{k+1})$.

Furthermore, $r_{k+1,k+1} = \| u_{k+1}'' \| \geq 0$, completing the induction step. $\square$

Theorem 12.1/12.4 For every real $n \times n$ matrix A, $\exists$ a sequence ... where each H_i is either a Householder

Theorem 12.1/12.4 For every real $n \times n$ matrix A, $\exists$ a sequence $H_1, \dots, H_n$ of matrices, where each H_i is either a Householder matrix or the identity, and an upper triangular matrix R, s.t.

$$R = H_n \cdots H_2 H_1 A.$$

Corollary: For any $A \in \mathbb{R}^{n \times n}$, $A = QR$, where Q is orthogonal and R is upper triangular with nonnegative diagonal entries.

Notes: (1) If A is invertible, and the diagonal of R is strictly positive, then $A = QR$ is unique.

(2) $\det(A) = (-1)^m \det(R)$.

(3) The condition number of A is preserved.

(4) The method also applies to any m-tuple of vectors $(v_1, \dots, v_n)$ with $m \leq n$. Then R is an upper triangular $m \times m$ matrix and Q is an $n \times m$ matrix with orthogonal columns $Q^T Q = I_m$.

(5) Similarly, applies to $m \times n$ matrix with $m \geq n$. Then $R \in \mathbb{R}^{n \times n}$ triangular and $Q \in \mathbb{R}^{m \times n}$ s.t. $Q^T Q = I_n$.

Aside: There is a complex analogue to QR decompositions, where Q is a unitary matrix