Towards QR factorization

We will now develop the *QR factorization*. We will show that certain matrices can be written as the product of matrices in special form.
Matrix factorizations are useful mathematically and computationally:

- *Mathematical:* They provide insight into the nature of matrices—each factorization gives us a new way to think about a matrix.
- *Computational:* They give us ways to compute solutions to fundamental computational problems involving matrices.

Matrices with mutually orthogonal columns

$$\begin{bmatrix} \mathbf{v}_1^{*\,T} \\ \hline \vdots \\ \hline \mathbf{v}_n^{*\,T} \end{bmatrix} \left[\begin{array}{c|c|c} \mathbf{v}_1^* & \cdots & \mathbf{v}_n^* \end{array}\right] = \begin{bmatrix} \|\mathbf{v}_1\|^2 & & \\ & \ddots & \\ & & \|\mathbf{v}_n\|^2 \end{bmatrix}$$

Cross-terms are zero because of mutual orthogonality.
To make the product into the identity matrix, can *normalize* the columns.

Normalizing a vector means scaling it to make its norm 1.

Just divide it by its norm.

```
>>> def normalize(v): return v/sqrt(v*v)
>>> q = normalize(list2vec[1,1,1])
>>> q * q
1.0000000000000002
>>> print(q)
     0     1     2
------------------
 0.577 0.577 0.577
```

Matrices with mutually orthogonal columns

$$\begin{bmatrix} \mathbf{v}_1^{*\,T} \\ \hline \vdots \\ \hline \mathbf{v}_n^{*\,T} \end{bmatrix} \left[\begin{array}{c|c|c} & & \\ \mathbf{v}_1^* & \cdots & \mathbf{v}_n^* \\ & & \end{array}\right] = \begin{bmatrix} \|\mathbf{v}_1\|^2 & & \\ & \ddots & \\ & & \|\mathbf{v}_n\|^2 \end{bmatrix}$$

Cross-terms are zero because of mutual orthogonality.

To make the product into the identity matrix, can *normalize* the columns.

Normalize columns $\left[\begin{array}{c|c|c} & & \\ \mathbf{v}_1^* & \cdots & \mathbf{v}_n^* \\ & & \end{array}\right] \Rightarrow \left[\begin{array}{c|c|c} & & \\ \mathbf{q}_1 & \cdots & \mathbf{q}_n \\ & & \end{array}\right]$

Matrices with mutually orthogonal columns

$$\begin{bmatrix} \mathbf{q}_1^T \\ \hline \vdots \\ \hline \mathbf{q}_n^T \end{bmatrix} \left[\begin{array}{c|c|c} \mathbf{q}_1 & \cdots & \mathbf{q}_n \end{array} \right] = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}$$

Normalize columns $\left[\begin{array}{c|c|c} \mathbf{v}_1^* & \cdots & \mathbf{v}_n^* \end{array} \right] \Rightarrow \left[\begin{array}{c|c|c} \mathbf{q}_1 & \cdots & \mathbf{q}_n \end{array} \right]$

Matrices with mutually orthogonal columns

$$\begin{bmatrix} \mathbf{q}_1^T \\ \hline \vdots \\ \hline \mathbf{q}_n^T \end{bmatrix} \left[\begin{array}{c|c|c} & & \\ \mathbf{q}_1 & \cdots & \mathbf{q}_n \\ & & \end{array}\right] = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}$$

Proposition: If columns of Q are mutually orthogonal with norm 1 then $Q^T Q$ is identity matrix.

Definition: Vectors that are mutually orthogonal and have norm 1 are *orthonormal*.

Definition: If columns of Q are orthonormal then we call Q a *column-orthogonal* matrix.
should be called *orthonormal* but oh well

Definition: If Q is square and column-orthogonal, we call Q an *orthogonal* matrix.

Proposition: If Q is an orthogonal matrix then its inverse is Q^T.

Projection onto columns of a column-orthogonal matrix

Suppose $\mathbf{q}_1, \ldots, \mathbf{q}_n$ are orthonormal vectors.

Projection of $\mathbf{b}$ onto $\mathbf{q}_j$ is $\mathbf{b}^{||\mathbf{q}_j} = \sigma_j\,\mathbf{q}_j$ where $\sigma_j = \frac{\langle \mathbf{q}_j, \mathbf{b}\rangle}{\langle \mathbf{q}_j, \mathbf{q}_j\rangle} = \langle \mathbf{q}_j, \mathbf{b}\rangle$

Vector $[\sigma_1, \ldots, \sigma_n]$ can be written using dot-product definition of matrix-vector multiplication:

$$\begin{bmatrix} \sigma_1 \\ \vdots \\ \sigma_n \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1 \cdot \mathbf{b} \\ \vdots \\ \mathbf{q}_n \cdot \mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1^T \\ \hline \vdots \\ \hline \mathbf{q}_n^T \end{bmatrix} \begin{bmatrix} \\ \mathbf{b} \\ \\ \end{bmatrix}$$

and linear combination $\sigma_1\,\mathbf{q}_1 + \cdots + \sigma_n\,\mathbf{q}_n = \left[\begin{array}{c|c|c} & & \\ \mathbf{q}_1 & \cdots & \mathbf{q}_n \\ & & \end{array}\right] \begin{bmatrix} \sigma_1 \\ \vdots \\ \sigma_n \end{bmatrix}$

Towards QR factorization

Orthogonalization of columns of matrix A gives us a representation of A as product of

- matrix with mutually orthogonal columns
- invertible triangular matrix

$$\begin{bmatrix} \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \cdots & \mathbf{v}_n \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \mathbf{v}_1^* & \mathbf{v}_2^* & \mathbf{v}_3^* & \cdots & \mathbf{v}_n^* \\ \\ \end{bmatrix} \begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} & & \alpha_{1n} \\ & 1 & \alpha_{23} & & \alpha_{2n} \\ & & 1 & & \alpha_{3n} \\ & & & \ddots & \\ & & & & \alpha_{n-1,n} \\ & & & & 1 \end{bmatrix}$$

Suppose columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent. Then $\mathbf{v}_1^*, \ldots, \mathbf{v}_n^*$ are nonzero.

- Normalize $\mathbf{v}_1^*, \ldots, \mathbf{v}_n^*$ (Matrix is called Q)
- To compensate, scale the rows of the triangular matrix. (Matrix is R)

The result is the QR factorization.

Q is a column-orthogonal matrix and R is an upper-triangular matrix.

Towards QR factorization

Orthogonalization of columns of matrix A gives us a representation of A as product of

- matrix with mutually orthogonal columns
- invertible triangular matrix

$$\left[\begin{array}{c|c|c|c|c} & & & & \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \cdots & \mathbf{v}_n \\ & & & & \end{array}\right] = \left[\begin{array}{c|c|c|c|c} & & & & \\ \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 & \cdots & \mathbf{q}_n \\ & & & & \end{array}\right] \begin{bmatrix} \|\mathbf{v}_1^*\| & \beta_{12} & \beta_{13} & & \beta_{1n} \\ & \|\mathbf{v}_2^*\| & \beta_{23} & & \beta_{2n} \\ & & \|\mathbf{v}_3^*\| & & \beta_{3n} \\ & & & \ddots & \\ & & & & \beta_{n-1,n} \\ & & & & \|\mathbf{v}_n^*\| \end{bmatrix}$$

Suppose columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent. Then $\mathbf{v}_1^*, \ldots, \mathbf{v}_n^*$ are nonzero.

- Normalize $\mathbf{v}_1^*, \ldots, \mathbf{v}_n^*$ (Matrix is called Q)
- To compensate, scale the rows of the triangular matrix. (Matrix is R)

The result is the QR factorization.

Q is a column-orthogonal matrix and R is an upper-triangular matrix.