Applied Numerical Linear Algebra. Lecture 9

Householder Transformations

A Householder transformation (or reflection) is a matrix of the form $P = I - 2uu^T$ where $\|u\|_2 = 1$. It is easy to see that $P = P^T$ and $P \cdot P^T = (I - 2uu^T)(I - 2uu^T) = I - 4uu^T + 4uu^T uu^T = I$, so P is a symmetric, orthogonal matrix. It is called a reflection because Px is reflection of x in the plane through 0 perpendicular to u.

Given a vector x, it is easy to find a Householder reflection $P = I - 2uu^T$ to zero out all but the first entry of x: $Px = [c, 0, \ldots, 0]^T = c \cdot e_1$. We do this as follows. Write $Px = (I - 2uu^T)x = x - 2u(u^T x) = c \cdot e_1$ so from that equation we get $u = \frac{1}{2(u^T x)}(x - ce_1)$, i.e., u is a linear combination of x and e_1. Since $\|x\|_2 = \|Px\|_2 = |c|$, u must be parallel to the vector $\tilde{u} = x \pm \|x\|_2 e_1$, and so $u = \tilde{u}/\|\tilde{u}\|_2$. One can verify that either choice of sign yields a u satisfying $Px = ce_1$, as long as $\tilde{u} \neq 0$. We will use $\tilde{u} = x + sign(x_1)e_1$, since this means that there is no cancellation in computing the first component of u. Here, x_k is to be the pivot coordinate in the vector x after which all entries are 0 in matrix A. In summary, we get

$$\tilde{u} = \begin{bmatrix} x_1 + sign(x_1) \cdot \|x\|_2 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \; with \; u = \frac{\tilde{u}}{\|\tilde{u}\|_2}.$$

We write this as $u = House(x)$. (In practice, we can store $\tilde{u}$ instead of u to save the work of computing u, and use the formula $P = I - (2/\|\tilde{u}\|_2^2)\tilde{u}\tilde{u}^T$ instead of $P = I - 2uu^T$.)

Idea of Householder transformation

We show how to compute the QR decomposition of a 5-by-4 matrix A using Householder transformations. This example will make the pattern for general m-by-n matrices evident. In the matrices below, P_i is an orthogonal matrix, x denotes a generic nonzero entry, and o denotes a zero entry.

1. Choose P_1 so

$$A_1 \equiv P_1 A = \begin{bmatrix} x & x & x & x \\ o & x & x & x \\ o & x & x & x \\ o & x & x & x \\ o & x & x & x \end{bmatrix}.$$

2. Choose $P_2 = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & P_2' \end{array}\right]$ so

$$A_2 \equiv P_2 A_1 = \begin{bmatrix} x & x & x & x \\ o & x & x & x \\ o & o & x & x \\ o & o & x & x \\ o & o & x & x \end{bmatrix}.$$

3. Choose $P_3 = \left[\begin{array}{cc|c} 1 & & \\ & & 0 \\ & 1 & \\ \hline \multicolumn{2}{c|}{0} & P_3' \end{array}\right]$ so

$$A_3 \equiv P_3 A_2 = \begin{bmatrix} x & x & x & x \\ o & x & x & x \\ o & o & x & x \\ o & o & o & x \\ o & o & o & x \end{bmatrix}.$$

4. Choose $P_4 = \left[\begin{array}{ccc|c} 1 & & & \\ & 1 & & 0 \\ & & 1 & \\ \hline & 0 & & P_4' \end{array}\right]$ so

$$\tilde{R} := A_4 \equiv P_4 A_3 = \begin{bmatrix} x & x & x & x \\ o & x & x & x \\ o & o & x & x \\ o & o & o & x \\ o & o & o & o \end{bmatrix}.$$

Here, we have chosen a Householder matrix P_i' to zero out the subdiagonal entries in column i; this does not disturb the zeros already introduced in previous columns.

Idea of Householder transformation

We observe that we have performed decomposition

$$A_4 = P_4 P_3 P_2 P_1 A. \tag{1}$$

Let us denote the final triangular matrix A_4 as $\tilde{R} \equiv A_4$. Then using (1) we observe that matrix A is obtained via decomposition

$$A = P_1^T P_2^T P_3^T P_4^T \tilde{R} = QR, \tag{2}$$

which is our desired QR decomposition. Here, the matrix Q is the first four columns of $P_1^T P_2^T P_3^T P_4^T = P_1 P_2 P_3 P_4$ (since all P_i are symmetric), and R is the first four rows of $\tilde{R}$.

$\diamond$

Here is the general algorithm for QR decomposition using Householder transformations.

ALGORITHM *QR factorization using Householder reflections:*

$$
\begin{aligned}
&\textit{for } i = 1 \textit{ to } \min(m-1, n) \\
&\quad u_i = House(A(i:m, i)) \\
&\quad P_i' = I - 2u_i u_i^T \\
&\quad A(i:m, i:n) = P_i' A(i:m, i:n) \\
&\textit{end for}
\end{aligned}
$$

QR decomposition using Householder reflections

We can use Householder reflections to calculate the QR factorization of an m-by-n matrix A with $m \geq n$.

- Let $\mathbf{x}$ be an arbitrary real m-dimensional column vector of A such that $||\mathbf{x}|| = |\alpha|$ for a scalar α.
- If the algorithm is implemented using floating-point arithmetic, then α should get the opposite sign as the k-th coordinate of $\mathbf{x}$, where x_k is to be the pivot coordinate after which all entries are 0 in matrix A's final upper triangular form, to avoid loss of significance.

Then, where $\mathbf{e}_1$ is the vector $(1, 0, ..., 0)^T$, $|| \cdot ||$ is the Euclidean norm and I is an m-by-m identity matrix, set

$$\begin{aligned}
\mathbf{v} &= \mathbf{x} + \alpha \mathbf{e}_1, \\
\alpha &= -sign(x_1)||\mathbf{x}||, \\
\mathbf{u} &= \frac{\mathbf{v}}{\|\mathbf{v}\|}, \\
Q &= I - 2\mathbf{u}\mathbf{u}^T.
\end{aligned}$$

In the case of complex A set

$$Q = I - (1 + w)\mathbf{u}\mathbf{u}^H,$$

where $w = \mathbf{x}^H\mathbf{u}/\mathbf{u}^H\mathbf{x}$ and where $\mathbf{x}^H$ is the conjugate transpose (transjugate) of $\mathbf{x}$,
Q is an m-by-m Householder matrix and

$$Q\mathbf{x} = (\alpha, 0, \cdots, 0)^T.$$

QR decomposition using Householder reflections.

Example

Let us calculate the decomposition of

$$A = \begin{pmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{pmatrix}.$$

First, we need to find a reflection that transforms the first column of matrix A, vector $\mathbf{x} = \mathbf{a}_1 = (12, 6, -4)^T$, to
$\|\mathbf{x}\| \mathrm{e}_1 = \|\mathbf{a}_1\| \mathrm{e}_1 = (14, 0, 0)^T$.
Now,

$$\mathbf{v} = \mathbf{x} + \alpha \mathbf{e}_1,$$

where

$$\alpha = -sign(x_1)||\mathbf{x}||,$$

and

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|}.$$

Example

Here, $||x|| = \sqrt{12^2 + 6^2 + (-4)^2} = 14$,

$$\alpha = -sign(12)||\mathbf{x}|| = -14 \text{ for } \mathbf{x} = \mathbf{a}_1 = (12, 6, -4)^T$$

Therefore

$$\mathbf{v} = \mathbf{x} + \alpha \mathbf{e}_1 = (-2, 6, -4)^T = (2)(-1, 3, -2)^T$$

and $\mathbf{u} = \frac{\mathbf{v}}{||\mathbf{v}||} = \frac{1}{\sqrt{14}}(-1, 3, -2)^T$, and then

$$Q_1 = I - \frac{2}{\sqrt{14}\sqrt{14}} \begin{pmatrix} -1 \\ 3 \\ -2 \end{pmatrix} \begin{pmatrix} -1 & 3 & -2 \end{pmatrix}$$

$$= I - \frac{1}{7} \begin{pmatrix} 1 & -3 & 2 \\ -3 & 9 & -6 \\ 2 & -6 & 4 \end{pmatrix}$$

$$= \begin{pmatrix} 6/7 & 3/7 & -2/7 \\ 3/7 & -2/7 & 6/7 \\ -2/7 & 6/7 & 3/7 \end{pmatrix}.$$

Example

Now observe:

$$A_1 = Q_1 A = \begin{pmatrix} 14 & 21 & -14 \\ 0 & -49 & -14 \\ 0 & 168 & -77 \end{pmatrix},$$

so we already have almost a triangular matrix. We only need to zero the (3, 2) entry.
Take the $(1,1)$ minor, and then apply the process again to

$$A' = M_{11} = \begin{pmatrix} -49 & -14 \\ 168 & -77 \end{pmatrix}.$$

By the same method as above we first need to find a reflection that transforms the first column of matrix A', vector $\mathbf{x} = (-49, 168)^T$, to $\|\mathbf{x}\| \, \mathrm{e}_1 = (175, 0)^T$.

Example

Here, $||x|| = \sqrt{(-49)^2 + 168^2} = 175$,

$$\alpha = -sign(-49)||\mathbf{x}|| = 175 \text{ and } \mathbf{x} = (-49, 168)^T.$$

Therefore

$$\mathbf{v} = \mathbf{x} + \alpha\mathbf{e}_1 = (-49, 168)^T + (175, 0)^T = (126, 168)^T,$$

$||\mathbf{v}|| = \sqrt{126^2 + 168^2} = \sqrt{44100} = 210$ and
$\mathbf{u} = \frac{\mathbf{v}}{||\mathbf{v}||} = (126/210, 168/210)^T = (3/5, 4/5)^T$.

Example

$$Q_2' = I - 2\begin{pmatrix} 3/5 \\ 4/5 \end{pmatrix}\begin{pmatrix} 3/5 & 4/5 \end{pmatrix}$$

or

$$\begin{aligned} Q_2' &= I - 2\begin{pmatrix} 9/25 & 12/25 \\ 12/25 & 16/25 \end{pmatrix} \\ &= \begin{pmatrix} 7/25 & -24/25 \\ -24/25 & -7/25 \end{pmatrix} \end{aligned}$$

Finally, we obtain the matrix of the Householder transformation Q_2 such that

$$Q_2 = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_2', \end{array}\right]$$

to get

$$Q_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 7/25 & -24/25 \\ 0 & -24/25 & -7/25. \end{pmatrix}$$

Example

Now, we have obtained $Q_2 A_1 = R$ which will be upper triangular matrix R. Thus, $R = Q_1 A_1$ and the matrix Q in QR decomposition of A can be obtained as follows:

$$
\begin{aligned}
Q_2 Q_1 A &= R, \\
Q_1^T Q_2^T Q_2 Q_1 A &= Q_1^T Q_2^T R \\
A = Q_1^T Q_2^T R &= QR, \\
\text{with} Q &= Q_1^T Q_2^T.
\end{aligned}
$$

$$
Q = Q_1^T Q_2^T = \begin{pmatrix} 6/7 & 69/175 & -58/175 \\ 3/7 & -158/175 & 6/175 \\ -2/7 & -6/35 & -33/35 \end{pmatrix}
$$

Example

Then

$$Q = Q_1^T Q_2^T = \begin{pmatrix} 0.8571 & 0.3943 & -0.3314 \\ 0.4286 & -0.9029 & 0.0343 \\ -0.2857 & -0.1714 & -0.9429 \end{pmatrix}$$

$$R = Q_2 A_1 = Q_2 Q_1 A = Q^T A = \begin{pmatrix} 14 & 21 & -14 \\ 0 & -175 & 70 \\ 0 & 0 & 35 \end{pmatrix}.$$

The matrix Q is orthogonal and R is upper triangular, so $A = QR$.

Example

Compute QR decomposition of the matrix A using Householder reflections:

$$\mathbf{A} = \begin{bmatrix} 4 & 4 & 3 \\ 0 & 3 & 1 \\ 3 & 4 & 7 \end{bmatrix}$$

Example

First, we need to find a reflection that transforms the first column of matrix A

$$\mathbf{A} = \begin{bmatrix} 4 & 4 & 3 \\ 0 & 3 & 1 \\ 3 & 4 & 7 \end{bmatrix}$$

We have:

$$\mathbf{u} = \mathbf{x} + \alpha \mathbf{e}_1,$$

where $\mathbf{x} = (4, 0, 3)^T$, $\alpha = -sign(4) \cdot ||x||$

$$\mathbf{v} = \frac{\mathbf{u}}{\|\mathbf{u}\|}.$$

Here,

$$\alpha = -5.$$

Therefore

$$\mathbf{u} = (-1, 0, 3)^T, \quad ||u|| = \sqrt{10}.$$

and $\mathbf{v} = \frac{1}{\sqrt{10}}(-1, 0, 3)^T$, and then

Example

$$P_1 = I - \frac{2}{\sqrt{10}\sqrt{10}} \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix} \begin{pmatrix} -1 & 0 & 3 \end{pmatrix}$$
$$= I - \frac{1}{5} \begin{pmatrix} 1 & 0 & -3 \\ 0 & 0 & 0 \\ -3 & 0 & 9 \end{pmatrix}$$
$$= \begin{pmatrix} 4/5 & 0 & 3/5 \\ 0 & 1 & 0 \\ 3/5 & 0 & -4/5 \end{pmatrix}.$$

Now observe:

$$P_1 A = \begin{pmatrix} 5 & 5.6 & 6.6 \\ 0 & 3 & 1 \\ 0 & -0.8 & -3.8 \end{pmatrix},$$

so we already have almost a triangular matrix. We only need to zero the (3, 2) entry.

Example

Take the $(1, 1)$ minor, and then apply the process again to

$$A' = M_{11} = \begin{pmatrix} 3 & 1 \\ -0.8 & -3.8 \end{pmatrix}.$$

We have:

$$\mathbf{u} = \mathbf{x} + \alpha \mathbf{e}_1,$$

where $\mathbf{x} = (3, -0.8)^T$, $\alpha = -sign(3) \cdot ||x||$

$$\mathbf{v} = \frac{\mathbf{u}}{\|\mathbf{u}\|}.$$

Here,

$$\alpha = -3.1048.$$

Therefore

$$\mathbf{u} = (-0.1048, -0.8)^T, \quad ||u|| = 0.8068.$$

and $\mathbf{v} = \frac{1}{0.8068}(-0.1048, -0.8)^T$,

Example

and then

$$P_2' = I - \frac{2}{0.651}\begin{pmatrix}-0.1048\\-0.8\end{pmatrix}\begin{pmatrix}-0.1048 & -0.8\end{pmatrix}$$
$$= I - \frac{2}{0.651}\begin{pmatrix}0.011 & 0.0838\\0.0838 & 0.64\end{pmatrix}$$
$$= \begin{pmatrix}0.9662 & -0.2575\\0.2575 & -0.9662\end{pmatrix}.$$

Then the second matrix of the Householder transformation is

$$P_2 = \begin{pmatrix}1 & 0 & 0\\0 & 0.9662 & -0.2575\\0 & -0.2575 & -0.9662\end{pmatrix}$$

Now, we find

$$R = P_2 P_1 A = \begin{pmatrix}5 & 5.6 & 6.6\\0 & 3.1046 & 1.9447\\0 & 0.0005 & 3.4141\end{pmatrix}.$$

Tridiagonalization using Householder transformation

This procedure is taken from the book: Numerical Analysis, Burden and Faires, 8th Edition.
In the first step, to form the Householder matrix in each step we need to determine α and r, which are given by:

$$\alpha = -\text{sgn}(a_{21})\sqrt{\sum_{j=2}^{n} a_{j1}^2};$$

$$r = \sqrt{\frac{1}{2}(\alpha^2 - a_{21}\alpha)};$$

From α and r, construct vector v:

$$v^{(1)} = \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix},$$

where $v_1 = 0;, v_2 = \frac{a_{21}-\alpha}{2r}$, and

$$v_k = \frac{a_{k1}}{2r} \quad \text{for each} \quad k = 3, 4..n$$

Then compute:

$$P^{(1)} = I - 2v^{(1)}(v^{(1)})^T$$

and obtaing matrix $A^{(1)}$ as

$$A^{(1)} = P^{(1)}AP^{(1)}$$

Having found $P^{(1)}$ and computed $A^{(1)}$ the process is repeated for $k = 2, 3, ..., n$ as follows:

$$\alpha = -\text{sgn}(a_{k+1,k})\sqrt{\sum_{j=k+1}^{n} a_{jk}^2};$$

$$r = \sqrt{\frac{1}{2}(\alpha^2 - a_{k+1,k}\alpha)};$$

$$v_1^{(k)} = v_2^{(k)} = .. = v_k^{(k)} = 0;$$

$$v_{k+1}^{(k)} = \frac{a_{k+1,k} - \alpha}{2r}$$

$$v_j^{(k)} = \frac{a_{jk}}{2r} \quad \text{for} \quad j = k+2; k+3, ..., n$$

$$P^{(k)} = I - 2v^{(k)}(v^{(k)})^T$$

$$A^{(k+1)} = P^{(k)}A^{(k)}P^{(k)}$$

Example 1

Example

In this example, the given matrix A is transformed to the similar tridiagonal matrix A_1 by using Householder Method. We have

$$\mathbf{A} = \begin{bmatrix} 5 & 1 & 0 \\ 1 & 6 & 3 \\ 0 & 3 & 7 \end{bmatrix},$$

Example

Steps:

1. First compute α as

$$\alpha = -\text{sgn}(a_{21})\sqrt{\sum_{j=2}^{n} a_{j1}^2} = -\sqrt{(a_{21}^2 + a_{31}^2)} = -\sqrt{(1^2 + 0^2)} = -1.$$

2. Using α we find r as

$$r = \sqrt{\frac{1}{2}(\alpha^2 - a_{21}\alpha)} = \sqrt{\frac{1}{2}((-1)^2 - 1 \cdot (-1))} = 1.$$

Example

3. From α and r, construct vector v:

$$v^{(1)} = \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix},$$

where $v_1 = 0;, v_2 = \frac{a_{21} - \alpha}{2r}$, and

$$v_k = \frac{a_{k1}}{2r} \quad \text{for each} \quad k = 3, 4..n$$

Example

To do tridiagonal matrix we compute:

$$\begin{aligned} v_1 &= 0, \\ v_2 &= \frac{a_{21} - \alpha}{2r} = \frac{1 - (-1)}{2 \cdot 1} = 1, \\ v_3 &= \frac{a_{31}}{2r} = 0. \end{aligned}$$

and we have

$$v^{(1)} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

Example

Then compute matrix $P^{(1)}$

$$P^{(1)} = I - 2v^{(1)}(v^{(1)})^T$$

and

$$P^{(1)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

After that we can obtain matrix $A^{(1)}$ as

$$A^{(1)} = P^{(1)}AP^{(1)} = \begin{bmatrix} 5 & -1 & 0 \\ -1 & 6 & -3 \\ 0 & -3 & 7. \end{bmatrix}$$

Example 2

Example

In this example, the given matrix A is transformed to the similar tridiagonal matrix A_2 by using Householder Method. We have

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & -2 & 2 \\ 1 & 2 & 0 & 1 \\ -2 & 0 & 3 & -2 \\ 2 & 1 & -2 & -1 \end{bmatrix},$$

Example

Steps:

1. First compute α as

$$\alpha = -\text{sgn}(a_{21})\sqrt{\sum_{j=2}^{n} a_{j1}^2} = (-1) \cdot \sqrt{(a_{21}^2 + a_{31}^2 + a_{41}^2)}$$

$$= -1 \cdot (1^2 + (-2)^2 + 2^2) = (-1) \cdot \sqrt{1+4+4} = -\sqrt{9} = -3.$$

2. Using α we find r as

$$r = \sqrt{\frac{1}{2}(\alpha^2 - a_{21}\alpha)} = \sqrt{\frac{1}{2}((-3)^2 - 1 \cdot (-3))} = \sqrt{6}.$$

Example

3. From α and r, construct vector v:

$$v^{(1)} = \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix},$$

where $v_1 = 0;, v_2 = \frac{a_{21} - \alpha}{2r}$, and

$$v_k = \frac{a_{k1}}{2r} \quad \text{for each} \quad k = 3, 4..n$$

Example

To do that we compute:

$$v_1 = 0,$$
$$v_2 = \frac{a_{21} - \alpha}{2r} = \frac{1 - (-3)}{2 \cdot \sqrt{6}} = \frac{2}{\sqrt{6}}$$
$$v_3 = \frac{a_{31}}{2r} = \frac{-2}{2 \cdot \sqrt{6}} = \frac{-1}{\sqrt{6}}$$
$$v_4 = \frac{a_{41}}{2r} = \frac{2}{2 \cdot \sqrt{6}} = \frac{1}{\sqrt{6}}.$$

and we have

$$v^{(1)} = \begin{bmatrix} 0 \\ \frac{2}{\sqrt{6}} \\ \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix},$$

Example

Then compute matrix $P^{(1)}$

$$P^{(1)} = I - 2v^{(1)}(v^{(1)})^T = I - 2 \cdot \begin{bmatrix} 0 \\ \frac{2}{\sqrt{6}} \\ \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix} \cdot \begin{bmatrix} 0 & \frac{2}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

and

$$P^{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1/3 & 2/3 & -2/3 \\ 0 & 2/3 & 2/3 & 1/3 \\ 0 & -2/3 & 1/3 & 2/3 \end{bmatrix}$$

After that we can obtain matrix $A^{(1)}$ as

$$A^{(1)} = P^{(1)} A P^{(1)}$$

Example

Thus, the first Householder matrix:

$$P^{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1/3 & 2/3 & -2/3 \\ 0 & 2/3 & 2/3 & 1/3 \\ 0 & -2/3 & 1/3 & 2/3 \end{bmatrix},$$

$$A^{(1)} = P^{(1)} A P^{(1)} = \begin{bmatrix} 4 & -3 & 0 & 0 \\ -3 & 10/3 & 1 & 4/3 \\ 0 & 1 & 5/3 & -4/3 \\ 0 & 4/3 & -4/3 & -1 \end{bmatrix},$$

Example

Next, having found $A^{(1)}$ we need to construct $A^{(2)}$ and $P^{(2)}$. When $k = 2$ we have following formulas:

$$\alpha = -\mathrm{sgn}(a_{3,2})\sqrt{\sum_{j=3}^{4} a_{j2}^2} = -\mathrm{sgn}(1)\sqrt{a_{3,2}^2 + a_{4,2}^2} = -\sqrt{1 + \frac{16}{9}} = -\frac{5}{3};$$

$$r = \sqrt{\frac{1}{2}(\alpha^2 - a_{3,2}\cdot\alpha)} = \sqrt{\frac{20}{9}};$$

$$v_1^{(2)} = v_2^{(2)} = 0;$$

$$v_3^{(2)} = \frac{a_{3,2} - \alpha}{2r} = \frac{2}{\sqrt{5}}$$

$$v_4^{(2)} = \frac{a_{4,2}}{2r} = \frac{1}{\sqrt{5}}.$$

and thus new vector v will be: $v^{(2)} = (0, 0, \frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}})^T$ and the new Householder matrix $P^{(2)}$ will be

Example

$$P^{(2)} = I - 2v^{(2)}(v^{(2)})^T = I - 2\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 4/5 & 2/5 \\ 0 & 0 & 2/5 & 1/5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -3/5 & -4/5 \\ 0 & 0 & -4/5 & 3/5 \end{bmatrix}$$

and thus

$$A^{(2)} = P^{(2)}A^{(1)}P^{(2)} = \begin{bmatrix} 4 & -3 & 0 & 0 \\ -3 & 10/3 & -5/3 & 0 \\ 0 & -5/3 & -33/25 & 68/75 \\ 0 & 0 & 68/75 & 149/75 \end{bmatrix},$$

As we can see, the final result is a tridiagonal symmetric matrix which is similar to the original one. The process finished after 2 steps.

Given's Rotation

A Givens rotation is represented by a matrix of the form

$$G(i,j,\theta) = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & c & \cdots & -s & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & s & \cdots & c & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}$$

where $c = \cos(\theta)$ and $s = sin(\theta)$ appear at the intersections i-th and j-th rows and columns.

That is, the non-zero elements of Givens matrix is given by:

$$g_{k\,k} = 1 \qquad \text{for } k \neq i, j \tag{3}$$

$$g_{i\,i} = c \tag{4}$$

$$g_{j\,j} = c \tag{5}$$

$$g_{j\,i} = -s \tag{6}$$

$$g_{i\,j} = s \qquad \text{for } i > j \tag{7}$$

(sign of sine switches for $j > i$)

Given's Rotation

The product $G(i,j,\theta)x$ represents a counterclockwise rotation of the vector x in the (i,j) plane of θ radians, hence the name Givens rotation. When a Givens rotation matrix G multiplies another matrix, A, from the left, GA, only rows i and j of A are affected. Thus we restrict attention to the following problem. Given a and b, find $c = cos\theta$ and $s = sin\theta$ such that

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}.$$

Explicit calculation of θ is rarely necessary or desirable. Instead we directly seek c, s, and r. An obvious solution would be

$$r = \sqrt{a^2 + b^2} \tag{8}$$

$$c = a/r \tag{9}$$

$$s = -b/r. \tag{10}$$

Given's Rotation to get upper Triangular matrix

Example

Given the following 3x3 Matrix, perform two iterations of the Given's Rotation to bring the matrix to an upper Triangular matrix.

$$A = \begin{bmatrix} 6 & 5 & 0 \\ 5 & 1 & 4 \\ 0 & 4 & 3 \end{bmatrix}$$

In order to form the desired matrix, we must zero elements (2,1) and (3,2). We first select element (2,1) to zero. Using a rotation matrix of:

$$G_1 = \begin{bmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Example

We have the following matrix multiplication:

$$A_1 = G_1 \cdot A = \begin{bmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 6 & 5 & 0 \\ 5 & 1 & 4 \\ 0 & 4 & 3 \end{bmatrix} \tag{11}$$

Here, $a = 6, b = 5$ and we can compute r, c, s as:

$$\begin{aligned} r &= \sqrt{6^2 + 5^2} = 7.8102 \\ c &= 6/r = 0.7682 \\ s &= -5/r = -0.6402 \end{aligned}$$

Plugging in (11) computed values for c and s and performing the matrix multiplication (11) we get:

$$\mathrm{A}_1 = \begin{bmatrix} 7.8102 & 4.4813 & 2.5607 \\ 0 & -2.4327 & 3.0729 \\ 0 & 4 & 3 \end{bmatrix}$$

Example

We now want to zero element (3,2) to finish off the process. Using the same idea as before, we have a rotation matrix of:

$$G_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{bmatrix}$$

We have to do the following matrix multiplication:

$$A_2 = G_2 \cdot A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{bmatrix} \begin{bmatrix} 7.8102 & 4.4813 & 2.5607 \\ 0 & -2.4327 & 3.0729 \\ 0 & 4 & 3 \end{bmatrix} \tag{12}$$

with $a = -2.4327, b = 4$. Thus, we can compute new r, c, s:

$$r = \sqrt{(-2.4327)^2 + 4^2} = 4.6817 \tag{13}$$

$$c = -2.4327/r = -0.5196 \tag{14}$$

$$s = -4/r = -0.8544 \tag{15}$$

Example

Plugging in (12) these values for c and s and performing the multiplications gives us a new matrix:

$$R = A_2 = \begin{bmatrix} 7.8102 & 4.4813 & 2.5607 \\ 0 & 4.6817 & 0.9664 \\ 0 & 0 & -4.1843 \end{bmatrix}$$

Calculating the QR decomposition

Example

This new matrix R is the upper triangular matrix needed to perform an iteration of the QR decomposition. Q is now formed using the transpose of the rotation matrices in the following manner:

$Q = G_1^T G_2^T$

We note that

$G_2 G_1 A = R$

$G_1^T G_2^T G_2 G_1 A = G_1^T G_2^T R$

and thus

$A = G_1^T G_2^T R = QR$

with

$Q = G_1^T G_2^T$.

Performing this matrix multiplication yields:

$$Q = \begin{bmatrix} 0.7682 & 0.3327 & 0.5470 \\ 0.6402 & -0.3992 & -0.6564 \\ 0 & 0.8544 & -0.5196 \end{bmatrix}$$

Example

Obtain QR decomposition of the matrix A

$$\mathbf{A} = \begin{bmatrix} 4 & 4 & 3 \\ 3 & 3 & 1 \\ 0 & 4 & 7 \end{bmatrix}$$

using Given's rotation.
Hint:

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}.$$

We directly seek c, s, and r:

$$r = \sqrt{a^2 + b^2} \tag{16}$$

$$c = a/r \tag{17}$$

$$s = -b/r. \tag{18}$$

Example

To obtain QR decomposition of the matrix A

$$\mathbf{A} = \begin{bmatrix} 4 & 4 & 3 \\ 3 & 3 & 1 \\ 0 & 4 & 7 \end{bmatrix}$$

using Given's rotation we have to zero out $(2,1)$ and $(3,2)$ elements of the matrix A.

1. First, we zero out element $(2,1)$ of the matrix A.

To do that we compute c, s from the known $a = 4$ and $b = 3$ as

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}$$

to get:

$$\begin{aligned} r &= \sqrt{a^2 + b^2} = \sqrt{4^2 + 3^2} = 5, \\ c &= a/r = 0.8, \\ s &= -b/r = -0.6. \end{aligned}$$

Example

The first Given's matrix will be

$$\mathbf{G_1} = \begin{bmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

or

$$\mathbf{G_1} = \begin{bmatrix} 0.8 & 0.6 & 0 \\ -0.6 & 0.8 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Then

$$\mathbf{G_1} \cdot \mathbf{A} = \begin{bmatrix} 5 & 5 & 3 \\ 0 & 0 & -1 \\ 0 & 4 & 7 \end{bmatrix}$$

Example

2. Next step is to construct second Given's matrix G_2 in order to zero out $(3,2)$ element of the matrix $G_1 \cdot A$.
To do that we compute c, s from the known $a = 0$ and $b = 4$ as

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}$$

to get formulas:

$$r = \sqrt{a^2 + b^2} = \sqrt{0^2 + 4^2} = 4,$$
$$c = \frac{a}{r} = 0,$$
$$s = \frac{-b}{r} = -1.$$

Example

Thus, the second Given's matrix will be

$$\mathbf{G_2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{bmatrix}$$

or

$$\mathbf{G_2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

Example

Then upper triangular matrix R in the QR decomposition will be

$$\mathbf{R} = \mathbf{G_2} \cdot \mathbf{G_1} \cdot \mathbf{A} = \begin{bmatrix} 5 & 5 & 3 \\ 0 & 4 & 7 \\ 0 & 0 & 1 \end{bmatrix}$$

Then $A = G_1^T \cdot G_2^T \cdot R = QR$ will be QR decomposition of the matrix A with $Q = G_1^T \cdot G_2^T$ given by

$$\mathbf{Q} = \begin{bmatrix} 0.8 & 0 & 0.6 \\ 0.6 & 0 & -0.8 \\ 0 & 1 & 0 \end{bmatrix}$$

Example

We will construct a lower triangular matrix using Given's rotation from the matrix

$$A = \begin{bmatrix} 5 & 4 & 3 \\ 4 & 6 & 1 \\ 3 & 1 & 7 \end{bmatrix}.$$

Given's matrix for $j < k$

function [G] = GivensMatrixLow(A, j,k)

$a = A(k,k)$
$b = A(j,k)$
$r = sqrt(a^2 + b^2);$
$c = a/r;$
$s = -b/r;$
$G = eye(length(A));$
$G(j,j) = c;$
$G(k,k) = c;$
$G(j,k) = s;$
$G(k,j) = -s;$

>>G1up = GivensMatrixLow(A,2,3)

$$G1 = \begin{bmatrix} 1.000000000000000 & 0 & 0 \\ 0 & 0.989949493661166 & -0.141421356237310 \\ 0 & 0.141421356237310 & 0.989949493661166 \end{bmatrix}$$

>> A1 =G1*A

$$A1 = \begin{bmatrix} 5.000000000000000 & 4.000000000000000 & 3.000000000000000 \\ 3.535533905932737 & 5.798275605729690 & -0.000000000000000 \\ 3.535533905932738 & 1.838477631085023 & 7.071067811865475 \end{bmatrix}$$

>>G2 = GivensMatrixLow(A1,1,3)

$$G2 = \begin{bmatrix} 0.920574617898323 & 0 & -0.390566732942472 \\ 0 & 1.000000000000000 & 0 \\ 0.390566732942472 & 0 & 0.920574617898323 \end{bmatrix}$$

>> A2=G2*A1

$$A2 = \begin{bmatrix} 3.222011162644131 & 2.964250269632601 & -0.000000000000000 \\ 3.535533905932737 & 5.798275605729690 & -0.000000000000000 \\ 5.207556439232954 & 3.254722774520597 & 7.681145747868607 \end{bmatrix}$$

>>G3 = GivensMatrixLow(A2,1,2)

$$G3 = \begin{bmatrix} 0.890391914715406 & -0.455194725594918 & 0 \\ 0.455194725594918 & 0.890391914715406 & 0 \\ 0 & 0 & 1.000000000000000 \end{bmatrix}$$

>> A3=G3*A2

$$A3 = \begin{bmatrix} 1.259496302198541 & 0 & -0.000000000000000 \\ 4.614653291088246 & 6.512048806713364 & -0.000000000000000 \\ 5.207556439232954 & 3.254722774520597 & 7.681145747868607 \end{bmatrix}$$

Rank-deficient Least Squares Problems

Proposition

Let A be m by n with $m \geq n$ and rank $A = r < n$. Then there is an $n - r$ dimensional set of vectors that minimize $||Ax - b||_2$.

Proof

Let $Az = 0$. Then of x minimizes $||Ax - b||_2$ then $x + z$ also minimizes $||A(x + z) - b||_2$.

This means that the least-squares solution is not unique.

Moore-Penrose pseudoinverse for a full rank A

Definition

Suppose that A is m by n with $m > n$ and has full rank with $A = QR = U\Sigma V^T$ being a QR and SVD decompositions of A, respectively. Then

$$A^+ \equiv (A^T A)^{-1} A^T = R^{-1} Q^T = V \Sigma^{-1} U^T$$

is called the Moore-Penrose pseudoinverse of A. If $m < n$ then $A^+ \equiv A^T (AA^T)^{-1}$.

The pseudoinverse of A allows write solution of the full-rank overdetermined least squares problem as $x = A^+ b$. If A is square and a full rank then this formula reduces to $x = A^{-1} b$. The A^+ is computed as *pinv(A)* in Matlab.

$$A^+ \equiv (A^T A)^{-1} A^T = ((QR)^T QR)^{-1} (QR)^T = (R^T Q^T QR)^{-1} (QR)^T$$
$$= (R^T R)^{-1} R^T Q^T = R^{-1} Q^T;$$

$$A^+ \equiv (A^T A)^{-1} A^T = ((U\Sigma V^T)^T U\Sigma V^T)^{-1} \cdot (U\Sigma V^T)^T$$
$$= (V\Sigma U^T U\Sigma V^T)^{-1} V\Sigma U^T = (V\Sigma^2 V^T)^{-1} V\Sigma U^T = V\Sigma^{-1} U^T$$

Moore-Penrose pseudoinverse for rank-deficient A

Definition

Suppose that A is m by n with $m > n$ and is rank-deficient with rank $r < n$. Let $A = U\Sigma V^T = U_1\Sigma_1 V_1^T$ being a SVD decompositions of A such that

$$\mathrm{A} = [\mathrm{U}_1, U_2] \left[\begin{array}{c|c} \Sigma_1 & 0 \\ \hline 0 & 0 \end{array}\right] [V_1, V_2]^T = U_1\Sigma_1 V_1^T$$

Here, $size(\Sigma_1) = r \times r$ and is nonsingular, U_1 and V_1 have r columns. Then

$$A^+ \equiv V_1\Sigma_1^{-1}U_1^T$$

is called the Moore-Penrose pseudoinverse for rank-deficient A. The solution of the least-squares problem is always $x = A^+b$, when A is rank-deficient then x has minimum norm.

The next proposition states that if A is nearly rank deficient then the solution x of $Ax = b$ will be ill-conditioned and very large.

Proposition

Let $\sigma_{min} > 0$ is the smallest singular value of the nearly rank deficient A. Then

- 1. If x minimizes $||Ax - b||_2$, then $||x||_2 \geq \frac{|u_n^T b|}{\sigma_{min}}$ where u_n is the last column of U in SVD decomposition of $A = U\Sigma V^T$.
- 2. Changing b to $b + \delta b$ can change x to $x + \delta x$ where $||\delta x||_2$ can be estimated as $\frac{||\delta b||_2}{\sigma_{min}}$, or the solution is very ill-conditioned.

Proof

1: We have that for the case of full-rank matrix A the solution of $Ax = b$ is given by $x = (U\Sigma V^T)^{-1} b = V\Sigma^{-1} U^T b$. The matrix $A^+ = V\Sigma^{-1}U^T$ is Moore-Penrose pseodoinverse of A. Thus, we can write also this solution as $x = V\Sigma^{-1}U^T b = A^+ b$.

Then taking norms from both sides of above expression we have:

$$||x||_2 = ||\Sigma^{-1}U^T b||_2 \geq |(\Sigma^{-1}U^T b)_n| = \frac{|u_n^T b|}{\sigma_{min}}, \tag{19}$$

where $|(\Sigma^{-1}U^T b)_n|$ is the n-th column of this product.

2. We apply now (19) for $\|x + \delta x\|$ instead of $\|x\|$ to get:

$$\begin{aligned} ||x + \delta x||_2 &= ||\Sigma^{-1} U^T (b + \delta b)||_2 \geq |(\Sigma^{-1} U^T (b + \delta b))_n| \\ &= \frac{|u_n^T (b + \delta b)|}{\sigma_{min}} = \frac{|u_n^T b + u_n^T \delta b|}{\sigma_{min}}. \end{aligned} \tag{20}$$

We observe that $\frac{|u_n^T b|}{\sigma_{min}} + \frac{|u_n^T \delta b|}{\sigma_{min}} \leq ||x + \delta x||_2 \leq ||x||_2 + ||\delta x||_2$.
Choosing δb parallel to u_n and applying again (19) for estimation of $\|x\|_2$ we have

$$||\delta x||_2 \geq \frac{||\delta b||_2}{\sigma_{min}}. \tag{21}$$

In the next proposition we prove that the minimum norm solution x is unique and may be well-conditioned if the smallest nonzero singular value is not too small.

Proposition

When A is exactly singular, then x that minimize $||Ax - b||_2$ can be characterized as follows. Let $A = U\Sigma V^T$ have rank $r < n$. Write svd of A as

$$\text{A} = [\text{U}_1, U_2] \left[\begin{array}{c|c} \Sigma_1 & 0 \\ \hline 0 & 0 \end{array} \right] [V_1, V_2]^T = U_1 \Sigma_1 V_1^T$$

Here, $size(\Sigma_1) = r \times r$ and is nonsingular, U_1 and V_1 have r columns. Let $\sigma = \sigma_{min}(\Sigma_1). Then$

- 1. All solutions x can be written as $x = V_1 \Sigma_1^{-1} U_1^T + V_2 z$
- 2. The solution x has minimal norm $||x||_2$ when $z = 0$. Then $x = V_1 \Sigma_1^{-1} U_1^T$ and $||x||_2 \leq \frac{||b||_2}{\sigma}$.
- 3. Changing b to $b + \delta b$ can change x as $\frac{||\delta b||_2}{\sigma}$.

Proof

We choose the matrix $\tilde{U}$ such that $[U, \tilde{U}] = [U_1, U_2, \tilde{U}]$ be an $m \times m$ orthogonal matrix. Then

$$
\begin{aligned}
||Ax - b||_2^2 &= ||[U_1, U_2, \tilde{U}]^T (Ax - b)||_2^2 \\
&= \left|\left| \begin{bmatrix} \mathrm{U}_1^T \\ \mathrm{U}_2^T \\ \tilde{U}^T \end{bmatrix} (U_1 \Sigma_1 V_1^T x - b) \right|\right|_2^2 \\
&= ||[I^{r \times r}, O^{m \times (n-r)}, 0^{m \times m-n}]^T (\Sigma_1 V_1^T x - [U_1, U_2, \tilde{U}]^T \cdot b)||_2^2 \\
&= ||[\Sigma_1 V_1^T x - U_1^T b; -U_2^T b; -\tilde{U}^T b]^T||_2^2 \\
&= ||\Sigma_1 V_1^T x - U_1^T b||_2^2 + ||U_2^T b||_2^2 + ||\tilde{U}^T b||_2^2
\end{aligned}
$$

1. Then $||Ax - b||_2$ is minimized when $\Sigma_1 V_1^T x - U_1^T b = 0$. We can also write that the vector $x = (\Sigma_1 V_1^T)^{-1} U_1^T b + V_2 z$ or $x = V_1 \Sigma_1^{-1} U_1^T b + V_2 z$ is also solution of this minimization problem, because $V_1^T V_2 z = 0$ since columns of V_1 and V_2 are orthogonal.

2. Since columns of V_1 and V_2 are orthogonal, then by Pythagorean theorem we have that $\|x\|_2^2 = ||V_1 \Sigma_1^{-1} U_1^T b||^2 + ||V_2 z||^2$ which is minimized for $z = 0$.
3. Changing b to δb in the expression above we have:

$$||V_1 \Sigma_1^{-1} U_1^T \delta b||_2 \leq \|V_1 \Sigma_1^{-1} U_1^T\|_2 \cdot \|\delta b\|_2 = \|\Sigma_1^{-1}\|_2 \cdot \|\delta b\|_2 = \frac{||\delta b||_2}{\sigma}, \tag{22}$$

where σ is smallest nonzero singular value of A. In this proof we used properties of the norm: $\|QAZ\|_2 = \|A\|_2$ if Q, Z are orthogonal.

How to solve rank-deficient least squares problems using QR decomposition with pivoting

QR decomposition with pivoting is cheaper but can be less accurate than SVD technique for solution of rank-deficient least squares problems.
If A has a rank $r < n$ with independent r columns QR decomposition can look like that

$$A = QR = Q \cdot \begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \\ 0 & 0 \end{bmatrix}. \tag{23}$$

with nonzingular R_{11} is of the size $r \times r$ and R_{12} is of the size $r \times (n-r)$. We can try to get

$$R = \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \\ 0 & 0 \end{bmatrix}, \tag{24}$$

where elements of R_{22} are very small and are of the order $\varepsilon \|A\|_2$.

If we set $R_{22} = 0$ and choose $[Q, \tilde{Q}]$ which is square and orthogonal then we will minimize

$$\begin{aligned}
\|Ax - b\|_2^2 &= \left\| \begin{bmatrix} \mathrm{Q}^T \\ \tilde{Q}^T \end{bmatrix} (Ax - b) \right\|_2^2 \\
&= \left\| \begin{bmatrix} \mathrm{Q}^T \\ \tilde{Q}^T \end{bmatrix} (QRx - b) \right\|_2^2 \\
&= \left\| \begin{bmatrix} \mathrm{Rx - Q}^T b \\ \text{- } \tilde{Q}^T \text{ b} \end{bmatrix} \right\|_2^2 \\
&= \|Rx - Q^T b\|_2^2 + \|\tilde{Q}^T b\|_2^2 .
\end{aligned} \tag{25}$$

Here we again used properties of the norm: $\|QAZ\|_2 = \|A\|_2$ if Q, Z are orthogonal.

Let us now decompose $Q = [Q_1, Q_2]$ with $x = [x_1, x_2]^T$ and

$$R = \begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix} \tag{26}$$

such that equation (25) becomes

$$\begin{aligned} \|Ax - b\|_2^2 &= \left\| \begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathrm{x}_1 \\ \mathrm{x}_2 \end{bmatrix} - \begin{bmatrix} \mathrm{Q}_1^T b \\ \mathrm{Q}_2^T b \end{bmatrix} \right\|_2^2 + \|\tilde{Q}^T b\|_2^2 \\ &= \|R_{11}x_1 + R_{12}x_2 - Q_1^T b\|_2^2 + \|Q_2^T b\|_2^2 + \|\tilde{Q}^T b\|_2^2. \end{aligned} \tag{27}$$

We take now derivative with respect to x to get $(\|Ax - b\|_2^2)'_x = 0$. We see that minimum is achieved when

$$x = \begin{bmatrix} \mathrm{R}_{11}^{-1}(Q_1^T b - R_{12}x_2) \\ \mathrm{x}_2 \end{bmatrix} \tag{28}$$

for any vector x_2. If R_{11} is well-conditioned and $R_{11}^{-1}R_{12}$ is small than the choice $x_2 = 0$ will be good one.

The described method is not reliable for all rank-deficient least squares problems. This is because R can be nearly rank deficient for the case when no R_{22} is small. In this case can help QR decomposition with column pivoting: we factorize $AP = QR$ with permutation matrix P. To compute this permutation we do as follows:
1. In all columns from 1 to n at step i we select from the unfinished decomposition of part A in columns i to n and rows i to m the column with largest norm and exchange it with i-th column.
2. Then compute usual Householder transformation to zero out column i in entries $i+1$ to m.
Recent research is devoted to more advanced algorithms called rank-revealing QR algorithms which detects rank more faster and more efficient.
C. Bischof, Incremental condition estimation, *SIAM J.Matrix Anal.Appl.*, 11:312-322, 1990.
T.Chan, Rank revealing QR factorizations, *Linear Algebra Applications*, 88/89:67-82, 1987.

Nonsymmetric eigenvalue problems

- The algorithms for the eigenproblem can be divided into two groups: direct methods and iterative methods.
- We will consider only direct methods for computation of all eigenvalues and possibly, all eigenvectors (not iterative). However, we will still iterate. Typically used on dense matrices. Direct - since the method never fails to converge.
- Main direct method is QR iteration. No global convergence proof for this method.
- Iterative methods are applied to sparse matrices.
- Algorithms will involve transforming the matrix A into canonical forms. From these forms is easy to compute eigenvalues.

Canonical Forms

DEFINITION. The polynomial $p(\lambda) = det(A - \lambda I)$ is called the *characteristic polynomial of* A. The roots of $p(\lambda) = 0$ are the eigenvalues of A.
Since the degree of the characteristic polynomial $p(\lambda)$ equals n, the dimension of A, it has n roots, so A has n eigenvalues.

DEFINITION. A nonzero vector x satisfying $Ax = \lambda x$ is a *(right) eigenvector* for the eigenvalue λ. A nonzero vector y such that $y^*A = \lambda y^*$ is a *left eigenvector*. (Recall that $y^* = (\bar{y})^T$ is the *conjugate transpose* of y.)

DEFINITION. Let S be any nonsingular matrix. Then A and $B = S^{-1}AS$ are called *similar* matrices, and S is a similarity transformation.
PROPOSITION. Let $B = S^{-1}AS$, so A and B are similar. Then A and B have the same eigenvalues, and x (or y) is a right (or left) eigenvector of A if and only if $S^{-1}x$ (or S^*y) is a right (or left) eigenvector of B.

Proof. Using the fact that $\det(X \cdot Y) = \det(X) \cdot det(Y)$ for any square matrices X and Y, we can write

$$\det(A - \lambda I) = \det(S^{-1}(A - \lambda I)S) = \det(B - \lambda I).$$

So A and B have the same characteristic polynomials. $Ax = \lambda x$ holds if and only if $\underbrace{S^{-1}AS}_{B}\underbrace{S^{-1}x}_{x^*} = \lambda \underbrace{S^{-1}x}_{x^*}$ or $B(S^{-1}x) = \lambda(S^{-1}x)$. Similarly, $y^*A = \lambda y^*$ if and only if $y^*SS^{-1}AS = \lambda y^*S$ or $(S^*y)^*B = \lambda(S^*y)^*$. $\square$

THEOREM. Jordan canonical form. Given A, there exists a nonsingular S such that $S^{-1}AS = J$, where J is in *Jordan canonical form*. This means that J is block diagonal, with $J = diag(J_{n_1}(\lambda_1), J_{n_2}(\lambda_2), \ldots, J_{n_k}(\lambda_k))$ and

$$J_{n_i}(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{bmatrix}^{n_i \times n_i} .$$

J is unique, up to permutations of its diagonal blocks.

For a proof of this theorem, see a book on linear algebra such as [F. Gantmacher. The Theory of Matrices, vol. II (translation). Chelsea, New York, 1959] or [P. Halmos. Finite Dimensional Vector Spaces. Van Nostrand, New York, 1958].

- Each $J_m(\lambda)$ is called a *Jordan block* with eigenvalue λ of *algebraic multiplicity* m.
- If some $n_i = 1$, and λ_i is an eigenvalue of only that one Jordan block, then λ_i is called a *simple eigenvalue*.
- If all $n_i = 1$, so that J is diagonal, A is called *diagonalizable*; otherwise it is called *defective*.
- An n-by-n defective matrix does not have n eigenvectors. Although defective matrices are "rare" in a certain well-defined sense, the fact that some matrices do not have n eigenvectors is a fundamental fact confronting anyone designing algorithms to compute eigenvectors and eigenvalues.
- Symmetric matrices are never defective.

PROPOSITION.

- A Jordan block has one right eigenvector, $e_1 = [1, 0, \ldots, 0]^T$, and one left eigenvector, $e_n = [0, \ldots, 0, 1]^T$.
- Therefore, a matrix has n eigenvectors matching its n eigenvalues if and only if it is diagonalizable.
- In this case, $S^{-1}AS = diag(\lambda_i)$. This is equivalent to $AS = S\, diag(\lambda_i)$, so the i-th column of S is a right eigenvector for λ_i.
- It is also equivalent to $S^{-1}A = diag(\lambda_i)S^{-1}$, so the conjugate transpose of the ith row of S^{-1} is a left eigenvector for λ_i.
- If all n eigenvalues of a matrix A are distinct, then A is diagonalizable.

Proof. Let $J = J_m(\lambda)$ for ease of notation. It is easy to see $Je_1 = \lambda e_1$ and $e_n^T J = \lambda e_n^T$, so e_1 and e_n are right and left eigenvectors of J, respectively. To see that J has only one right eigenvector (up to scalar multiples), note that any eigenvector x must satisfy $(J - \lambda I)x = 0$, so x is in the null space of

$$J - \lambda I = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}.$$

But the null space of $J - \lambda I$ is clearly span(e_1), so there is just one eigenvector. If all eigenvalues of A are distinct, then all its Jordan blocks must be 1-by-1, so $J = diag(\lambda_1, \ldots, \lambda_n)$ is diagonal. □