Iterative Improvement of a Solution to Linear Equations

- When solving large sets of linear equations:
    - Not easy to obtain precision greater than computer’s limit.
    - Roundoff errors accumulation.
- There is a technique to recover the lost precision.

Iterative Improvement

- Suppose that a vector X is the exact solution of the linear set:

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

……. (1)

- Suppose after solving the linear set we get x with some errors (due to round offs) that is: $\mathbf{x} + \delta\mathbf{x}$

- Multiplying this solution by A will give us b with som

$$\mathbf{A} \cdot (\mathbf{x} + \delta\mathbf{x}) = \mathbf{b} + \delta\mathbf{b}$$

- …… •

(2)

Iterative Improvement

- Subtracting (1) from (2) gives:

$$\mathbf{A} \cdot \delta\mathbf{x} = \delta\mathbf{b} \qquad \text{......(3)}$$

- Substituting (2) into (3) gives

$$\mathbf{A} \cdot \delta\mathbf{x} = \mathbf{A} \cdot (\mathbf{x} + \delta\mathbf{x}) - \mathbf{b}$$

- All right-hand side is known and we to solve for δx .

Iterative Improvement

- LU decomposition is calculated already, so we can use it.
- After solving , we subtract δx from initial solution.
- these steps can be applied iteratively until the convergence accrued.

Example results

```
x[0]= -0.3694685803566218340598936720198253169655799865722
x[1]= 2.1470611163870776394446693302597850561141967773437
x[2]= 0.2468441555473033788281611577986041083931922912597
x[3]= -0.10502171013263031373874412111035780981183052062988
------------------------------
r[0]= 0.000000000000000034727755418821271741155275430324117
r[1]= -0.0000000000000000600017888996225005776099713062206027
r[2]= 0.00000000000000000422453342199012543598084058171131
r[3]= -0.0000000000000000063324668554277985330405739223625
------------------------------
x[0]= -0.36946858035662216712680105956678744405508041381836
x[1]= 2.14706111638707808353387918032240122556686401367188
x[2]= 0.2468441555473033233170099265407770872116088867187
x[3]= -0.1050217101326302443498050820380740333348512649536
```

Initial solution

Restored precisions

Improved solution

Singular Value Decomposition

SVD - Overview

A technique for handling matrices (sets of equations) that do not have an inverse. This includes square matrices whose determinant is zero and all rectangular matrices.

Common usages include computing the least-squares solutions, rank, range (column space), null space and pseudoinverse of a matrix.

SVD - Basics

The SVD of a *m-by-n* matrix **A** is given by the formula :

$$A = UWV^T$$

Where :

U is a *m-by-n* matrix of the orthonormal eigenvectors of $\mathbf{AA^T}$

$\mathbf{V^T}$ is the transpose of a *n-by-n* matrix containing the orthonormal eigenvectors of $\mathbf{A^TA}$

W is a *n-by-n* Diagonal matrix of the *singular values* which are the square roots of the eigenvalues of $\mathbf{A^TA}$

The Algorithm

Derivation of the SVD can be broken down into two major steps [2] :

1. Reduce the initial matrix to bidiagonal form using Householder transformations
2. Diagonalize the resulting matrix using QR transformations

$$\begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{bmatrix} \rightarrow \begin{bmatrix} x & x & & & \\ & x & x & & \\ & & x & x & \\ & & & x & x \\ & & & & x \end{bmatrix} \rightarrow \begin{bmatrix} x & & & & \\ & x & & & \\ & & x & & \\ & & & x & \\ & & & & x \end{bmatrix}$$

Initial Matrix	Bidiagonal Form	Diagonal Form

Householder Transformations

A Householder matrix is a defined as :

$H = I - 2ww^T$

Where w is a unit vector with $|w|^2 = 1$.

It ends up with the following properties :

$H = H^T$

$H^{-1} = H^T$

$H^2 = I$ (Identity Matrix)

If multiplied by another matrix, it results in a new matrix with zero'ed elements in a selected row / column based on the values chosen for w.

Applying Householder

To derive the bidiagonal matrix, we apply successive Householder matrices :

$$P_1 \underset{\mathrm{M}}{\begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{bmatrix}} \rightarrow \underset{\mathrm{M}_1}{\begin{bmatrix} x & x & x & x & x \\ & x & x & x & x \\ & x & x & x & x \\ & x & x & x & x \\ & x & x & x & x \end{bmatrix}} S_1 \rightarrow P_2 \underset{\mathrm{M}_2}{\begin{bmatrix} x & x & & & \\ & x & x & x & x \\ & x & x & x & x \\ & x & x & x & x \\ & x & x & x & x \end{bmatrix}} \rightarrow$$

$$\underset{\mathrm{M}_3}{\begin{bmatrix} x & x & & & \\ & x & x & x & x \\ & & x & x & x \\ & & x & x & x \\ & & x & x & x \end{bmatrix}} S_2 \rightarrow \ldots\ldots \rightarrow P_m \underset{\mathrm{M}_m}{\begin{bmatrix} x & x & & & \\ & x & x & x & x \\ & & x & x & x \\ & & & & x \\ & & & & x \end{bmatrix}} \rightarrow$$

$$\underset{\mathrm{M}_\mathrm{n}}{\begin{bmatrix} x & x & & & \\ & x & x & & \\ & & x & x & x \\ & & & x & x \\ & & & & x \end{bmatrix}} S_n \rightarrow \underset{\mathrm{B}}{\begin{bmatrix} x & x & & & \\ & x & x & & \\ & & x & x & \\ & & & x & x \\ & & & & x \end{bmatrix}}$$

Application con't

From here we see :

$P_1 M = M_1$

$M_1 S_1 = M_2$

$P_2 M_2 = M_3$

….

$M_N S_N = B$ [If $M > N$, then $P_M M_M = B$]

This can be re-written in terms of M :

$M = P_1^T M_1 = P_1^T M_2 S_1^T = P_1^T P_2^T M_3 S_1^T = \ldots = P_1^T \ldots P_M^T B S_N^T \ldots S_1^T = P_1 \ldots P_M B S_N \ldots S_1$

(Because $H^T = H$)

Householder Derivation

Now that we've seen how Householder matrices are used, how do we get one?
Going back to its definition : $H = I - 2ww^T$

Which is defined in terms of *w* - which is defined as

$$w = \frac{(x-y)}{\|x-y\|} \quad \text{and} \quad Hx = y \quad \text{and} \quad \|x\| = \|y\|$$

To make the Householder matrix useful, *w* must be derived from the column (or row) we want to transform.

This is accomplished by setting *x* to row / column to transform and y to desired pattern.

$\|s\|$ (Length

Householder Example

To derive P_1 for the given matrix M : $\begin{bmatrix} 4 & 3 & 0 & 2 \\ 2 & 1 & 2 & 1 \\ 4 & 4 & 0 & 3 \end{bmatrix}$

M

We would have : $x = \begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix}$ $y = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$ With : $\|x\| = \sqrt{4^2 + 2^2 + 4^2} = \sqrt{36} = 6$

$\|y\| = \sqrt{6^2 + 0^2 + 0^2} = \sqrt{36} = 6$

This leads to : $w = \dfrac{(x-y)}{\|x-y\|} = \dfrac{(4-6),(2-0),(4-0)}{\|(4-6),(2-0),(4-0)\|} = \dfrac{[-2,2,4]}{\sqrt{-2^2+2^2+4^2}} = \left[\dfrac{-2}{\sqrt{24}}, \dfrac{2}{\sqrt{24}}, \dfrac{4}{\sqrt{24}}\right]$

Simplifying : $w = \left[\dfrac{-1}{\sqrt{6}}, \dfrac{1}{\sqrt{6}}, \dfrac{2}{\sqrt{6}}\right]$

Then : $$P_1 = I - 2ww^T = I - 2\begin{bmatrix} \dfrac{-1}{\sqrt{6}} \\ \dfrac{1}{\sqrt{6}} \\ \dfrac{2}{\sqrt{6}} \end{bmatrix}\left[\dfrac{-1}{\sqrt{6}}, \dfrac{1}{\sqrt{6}}, \dfrac{2}{\sqrt{6}}\right] = I - 2\begin{bmatrix} \frac{1}{6} & -\frac{1}{6} & -\frac{1}{3} \\ -\frac{1}{6} & \frac{1}{6} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix} = I - \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} & -\frac{2}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{4}{3} \end{bmatrix}$$

Example con't

Finally :

$$P_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} & -\frac{2}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{4}{3} \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \end{bmatrix}$$

With that :

$$M_1 = P_1 M = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} 4 & 3 & 0 & 2 \\ 2 & 1 & 2 & 1 \\ 4 & 4 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 6 & 5 & \frac{2}{3} & \frac{11}{3} \\ 0 & -1 & \frac{4}{3} & -\frac{2}{3} \\ 0 & 0 & -\frac{4}{3} & -\frac{1}{3} \end{bmatrix}$$

Which we can see zero'ed the first column.

P_1 can be verified by performing the reverse operation :

$$M = P_1 M_1 = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} 6 & 5 & \frac{2}{3} & \frac{11}{3} \\ 0 & -1 & \frac{4}{3} & -\frac{2}{3} \\ 0 & 0 & -\frac{4}{3} & -\frac{1}{3} \end{bmatrix} = \begin{bmatrix} 4 & 3 & 0 & 2 \\ 2 & 1 & 2 & 1 \\ 4 & 4 & 0 & 3 \end{bmatrix}$$

Example con't

Likewise the calculation of S_1 for $M_1 = \begin{bmatrix} 6 & 5 & \frac{2}{3} & \frac{11}{3} \\ 0 & -1 & \frac{4}{3} & -\frac{2}{3} \\ 0 & 0 & -\frac{4}{3} & -\frac{1}{3} \end{bmatrix}$

Would have : $x = \begin{bmatrix} 6 \\ 5 \\ \frac{2}{3} \\ \frac{11}{3} \end{bmatrix} \qquad y = \begin{bmatrix} \frac{6\sqrt{350}}{3} \\ 0 \\ 0 \end{bmatrix}$

With : $\|x\| = \sqrt{6^2 + 5^2 + \left(\frac{2}{3}\right)^2 + \left(\frac{11}{3}\right)^2} = \sqrt{\frac{674}{9}} \approx 8.6538366$

$$\|y\| = \sqrt{6^2 + \left(\frac{\sqrt{350}}{3}\right)^2 + 0^2 + 0} = \sqrt{\frac{674}{9}} \approx 8.6538366$$

This leads to : $w = \frac{(x-y)}{\|x-y\|} = [0,-0.314814,0.169789,0.933843]$

Then : $S_1 = I - 2ww^T = I - 2\begin{bmatrix} 0 \\ -0.314814 \\ 0.169789 \\ 0.933843 \end{bmatrix}[0,-0.314814,0.169789,0.933843]$

Example con't

Finally :

$$S_1 = I - 2\begin{bmatrix} 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.099108 & -0.053452 & -0.293987 \\ 0.0 & -0.053452 & 0.028828 & 0.158556 \\ 0.0 & -0.293987 & 0.158556 & 0.872063 \end{bmatrix} = \begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.801784 & 0.106904 & 0.587974 \\ 0.0 & 0.106904 & 0.942344 & -0.317112 \\ 0.0 & 0.587974 & -0.317112 & -0.744126 \end{bmatrix}$$

With that :

$$M_2 = M_1 S_1 = \begin{bmatrix} 6 & 5 & \frac{2}{3} & \frac{11}{3} \\ 0 & -1 & \frac{4}{3} & -\frac{2}{3} \\ 0 & 0 & -\frac{4}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.801784 & 0.106904 & 0.587974 \\ 0.0 & 0.106904 & 0.942344 & -0.317112 \\ 0.0 & 0.587974 & -0.317112 & -0.744126 \end{bmatrix}$$

$$\approx \begin{bmatrix} 6 & 5 & 0.0 & 0.0 \\ 0 & -1.05 & 1.36 & -0.51 \\ 0 & -0.34 & -1.15 & 0.67 \end{bmatrix}$$

Which we can see zero'ed the first row.

The QR Algorithm

As seen, the initial matrix is placed into bidiagonal form which results in the following decomposition :

$M = PBS$ with $P = P_1 \ldots P_N$ and $S = S_N \ldots S_1$

The next step takes B and converts it to the final diagonal form using successive QR transformations.

QR Decompositions

The QR decomposition is defined as :

$$M = QR$$

Where Q is an orthogonal matrix (such that $Q^T = Q^{-1}$, $Q^TQ = QQ^T = I$)

And R is an upper triangular matrix :

$$R = \begin{bmatrix} x & x & x & x \\ & x & x & x \\ & & x & x \\ & & & x \end{bmatrix}$$

It has the property such that $RQ = M_1$ to which another decomposition can be performed. Hence $M_1 = Q_1R_1$, $R_1Q_1 = M_2$ and so on. In practice, after enough decompositions, M_x will converge to the desired SVD diagonal matrix – W.

QR Decomposition con't

Because Q is orthogonal (meaning $QQ^T = Q^TQ = 1$), we can redefine M_x in terms of Q_{x-1} and M_{x-1} only :

$$R_{x-1}Q_{x-1} = M_x \rightarrow Q_{x-1}R_{x-1}Q_{x-1} = Q_{x-1}M_x \rightarrow Q_{x-1}^TQ_{x-1}R_{x-1}Q_{x-1} = M_x \rightarrow Q_{x-1}^TM_{x-1}Q_{x-1} = M_x$$

Which can be written as : $M_{x-1} = Q_{x-1}^TM_xQ_{x-1}$

Starting with $M_0 = M$, we can describe the entire decomposition of W as :

$$M_0 = Q_0^TM_1Q_0 = Q_0^TQ_1^TM_2Q_1Q_0 = \ldots = Q_0^TQ_1^T \ldots Q_w^TWQ_w \ldots Q_1Q_0$$

One question remains – How do we derive Q?

Multiple methods exist for QR decompositions – including Householder Transformations, Hessenberg Transformations, Given's Rotations, Jacobi Transformations, etc.

Unfortunately the algorithm from book is not explicit on its chosen methodology – possibly Givens as it is used by reference material.

QR Decomposition using Givens rotations

A Givens rotation is used to rotate a plane about two coordinates axes and can be used to zero elements similar to the householder reflection.

It is represented by a matrix of the form :

$$G(i,j,\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c_{ii} & s_{ji} & 0 \\ 0 & -s_{ij} & c_{jj} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$c = \cos(\theta)$$

$$s = \sin(\theta)$$

The multiplication $G^T A$* effects only the rows i and j in A.

Likewise the multiplication AG only effects the columns i and j.

[1] Shows transpose on pre-multiply – but examples do not appear to be transposed (i.e. –s is still located i,j).

Givens rotation

The zeroing of an element is performed by computing the c and s in the following system.

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sqrt{a^2+b^2} \\ 0 \end{bmatrix}$$

Where b is the element being zeroed and a is next to b in the preceding column / row.

This is results in : $c = \dfrac{a}{\sqrt{a^2+b^2}}$ $\quad s = \dfrac{b}{\sqrt{a^2+b^2}}$

matrix

The application of Givens rotations on a bidiagonal matrix looks like the following and results in its implicit QR decomposition.

$$\underset{B_1}{\begin{bmatrix} x & x & & & \\ & x & x & & \\ & & x & x & \\ & & & x & x \\ & & & & x \end{bmatrix}} J_1 \longrightarrow J_2 \underset{B_2}{\begin{bmatrix} x & x & & & \\ + & x & x & & \\ & & x & x & \\ & & & x & x \\ & & & & x \end{bmatrix}} \longrightarrow \underset{B_3}{\begin{bmatrix} x & x & + & & \\ & x & x & & \\ & & x & x & \\ & & & x & x \\ & & & & x \end{bmatrix}} J_3 \longrightarrow$$

$$J_4 \underset{B_4}{\begin{bmatrix} x & x & & & \\ & x & x & & \\ & + & x & x & \\ & & & x & x \\ & & & & x \end{bmatrix}} \longrightarrow \underset{B_5}{\begin{bmatrix} x & x & & & \\ & x & x & + & \\ & & x & x & \\ & & & x & x \\ & & & & x \end{bmatrix}} J_5 \longrightarrow J_6 \underset{B_6}{\begin{bmatrix} x & x & & & \\ & x & x & & \\ & & x & x & \\ & & + & x & x \\ & & & & x \end{bmatrix}} \longrightarrow$$

$$\underset{B_7}{\begin{bmatrix} x & x & & & \\ & x & x & & \\ & & x & x & + \\ & & & x & x \\ & & & & x \end{bmatrix}} J_7 \longrightarrow J_8 \underset{B_8}{\begin{bmatrix} x & x & & & \\ & x & x & & \\ & & x & x & \\ & & & x & x \\ & & & + & x \end{bmatrix}} \longrightarrow \underset{B`}{\begin{bmatrix} x & x & & & \\ & x & x & & \\ & & x & x & \\ & & & x & x \\ & & & & x \end{bmatrix}}$$

Givens and Bidiagonal

With the exception of J_1, J_x is the Givens matrix computed from the element being zeroed.

J_1 is computed from the following : $\begin{bmatrix} c & s \\ -s & s \end{bmatrix} \begin{bmatrix} d_1^2 - \lambda \\ d_1 f_1 \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix}$

Which is derived from B and the smallest eigenvalue (λ) of T

$$B = \begin{bmatrix} d_1 & f_1 & & & \\ & d_2 & f_2 & & \\ & & \cdots & f_{n-2} & \\ & & & d_{n-1} & f_{n-1} \\ & & & & d_n \end{bmatrix} \qquad T = \begin{bmatrix} d_{n-1}^2 + f_{n-2}^2 & d_{n-1} f_{n-1} \\ d_{n-1} f_{n-1} & d_n^2 + f_{n-1}^2 \end{bmatrix}$$

Bidiagonal and QR

This computation of J_1 causes the implicit formation of B^TB which causes :

$$B` = J_i \ldots J_4 J_2 B J_1 J_3 \ldots J_k = U_i \ldots U_4 U_2 B^T B J_1 J_3 \ldots J_k \approx Q^T B^T B Q$$

QR Decomposition Given's rotation example

Let $\mathbf{A}^{(0)} = \mathbf{A} = \begin{pmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ & 4 & 2 \\ & -1 & 0 \end{pmatrix}.$

1. Use a_{31} to eliminate a_{41}. $r_{3,4} = \sqrt{1^2 + 1^2} = \sqrt{2}$. $\begin{cases} \cos\theta_{3,4} = a_{31}/r = 1/\sqrt{2}, \\ \sin\theta_{3,4} = a_{41}/r = 1/\sqrt{2}. \end{cases}$

$$\mathbf{G}_{3,4}^{(1)} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \cos\theta_{3,4} & \sin\theta_{3,4} \\ & & -\sin\theta_{3,4} & \cos\theta_{3,4} \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1/\sqrt{2} & 1/\sqrt{2} \\ & & -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

$$\mathbf{A}^{(1)} = \mathbf{G}_{3,4}^{(1)}\mathbf{A}^{(0)} = \begin{pmatrix} 1 & -1 & 4 \\ & 4 & -2 \\ & 3/\sqrt{2} & \sqrt{2} \\ 0 & -5/\sqrt{2} & -\sqrt{2} \end{pmatrix}.$$

example

2. Use a_{21} to eliminate a_{31}. $r_{2,3} = \sqrt{1^2 + \sqrt{2}^2} = \sqrt{3}$. $\begin{cases} \cos\theta_{2,3} = a_{21}/r = \sqrt{2}/\sqrt{3}, \\ \sin\theta_{2,3} = a_{31}/r = 1/\sqrt{3}. \end{cases}$

$$\mathbf{G}_{2,3}^{(1)} = \begin{pmatrix} 1 & & & \\ & \cos\theta_{2,3} & \sin\theta_{2,3} & \\ & -\sin\theta_{2,3} & \cos\theta_{2,3} & \\ & & & 1 \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & 1/\sqrt{3} & \sqrt{2}/\sqrt{3} & \\ & -\sqrt{2}/\sqrt{3} & 1/\sqrt{3} & \\ & & & 1 \end{pmatrix}$$

$$\begin{aligned} \mathbf{A}^{(2)} &= \mathbf{G}_{2,3}^{(1)}\mathbf{A}^{(1)} = \begin{pmatrix} 1 & & & \\ & 1/\sqrt{3} & \sqrt{2}/\sqrt{3} & \\ & -\sqrt{2}/\sqrt{3} & 1/\sqrt{3} & \\ & & & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ \sqrt{2} & 3/\sqrt{2} & \sqrt{2} \\ 0 & -5/\sqrt{2} & -\sqrt{2} \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1 & 4 \\ \sqrt{3} & 7/\sqrt{3} & 0 \\ 0 & -5/\sqrt{6} & \sqrt{6} \\ 0 & -5/\sqrt{2} & -\sqrt{2} \end{pmatrix} \end{aligned}$$

QR Decomposition **Given's rotation**

3. Use a_{11} to eliminate a_{21}. $r_{1,2} = \sqrt{1^2 + \sqrt{3}^2} = 2$. $\begin{cases} \cos\theta_{1,2} = a_{11}/r = 1/2, \\ \sin\theta_{1,2} = a_{21}/r = \sqrt{3}/2. \end{cases}$

$$\mathbf{G}_{1,2}^{(1)} = \begin{pmatrix} \cos\theta_{1,2} & \sin\theta_{1,2} & & \\ -\sin\theta_{1,2} & \cos\theta_{1,2} & & \\ & & 1 & \\ & & & 1 \end{pmatrix} = \begin{pmatrix} 1/2 & \sqrt{3}/2 & & \\ -\sqrt{3}/2 & 1/2 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

$$\begin{aligned} \mathbf{A}^{(3)} &= \mathbf{G}_{1,2}^{(1)}\mathbf{A}^{(2)} = \begin{pmatrix} 1/2 & \sqrt{3}/2 & & \\ -\sqrt{3}/2 & 1/2 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 4 \\ \sqrt{3} & 7/\sqrt{3} & 0 \\ 0 & -5/\sqrt{6} & \sqrt{6} \\ 0 & -5/\sqrt{2} & -\sqrt{2} \end{pmatrix} \\ &= \begin{pmatrix} 2 & 3 & 2 \\ 0 & 5/\sqrt{3} & -2\sqrt{3} \\ 0 & -5/\sqrt{6} & \sqrt{6} \\ 0 & -5/\sqrt{2} & -\sqrt{2} \end{pmatrix} \end{aligned}$$

example

4. Use a_{32} to eliminate a_{42}. $r_{3,4} = \sqrt{(-5/\sqrt{6})^2 + (-5/\sqrt{2})^2} = 10/\sqrt{6}$.

$$\begin{cases} \cos\theta_{3,4} = a_{32}/r = -1/2, \\ \sin\theta_{3,4} = a_{42}/r = -\sqrt{3}/2. \end{cases}$$

$$\mathbf{G}_{3,4}^{(2)} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \cos\theta_{3,4} & \sin\theta_{3,4} \\ & & -\sin\theta_{3,4} & \cos\theta_{3,4} \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1/2 & -\sqrt{3}/2 \\ & & \sqrt{3}/2 & -1/2 \end{pmatrix}$$

$$\begin{aligned} \mathbf{A}^{(4)} &= \mathbf{G}_{3,4}^{(2)}\mathbf{A}^{(3)} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1/2 & -\sqrt{3}/2 \\ & & \sqrt{3}/2 & -1/2 \end{pmatrix} \begin{pmatrix} 2 & 3 & 2 \\ 0 & 5/\sqrt{3} & -2\sqrt{3} \\ 0 & -5/\sqrt{6} & \sqrt{6} \\ 0 & -5/\sqrt{2} & -\sqrt{2} \end{pmatrix} \\ &= \begin{pmatrix} 2 & 3 & 4 \\ 0 & 5/\sqrt{3} & -2\sqrt{3} \\ 0 & 10/\sqrt{6} & 0 \\ 0 & 0 & 2\sqrt{2} \end{pmatrix} \end{aligned}$$

Ref: An example of QR Decomposition, Che-Rung Lee, November

QR Decomposition Given's rotation example

5. Use a_{22} to eliminate a_{32}. $r_{2,3} = \sqrt{(10/\sqrt{6})^2 + (5/\sqrt{3})^2} = 5.$ $\begin{cases} \cos\theta_{2,3} = a_{22}/r = 1/\sqrt{3}, \\ \sin\theta_{2,3} = a_{32}/r = 2/\sqrt{6}. \end{cases}$

$$\mathbf{G}_{2,3}^{(2)} = \begin{pmatrix} 1 & & & \\ & \cos\theta_{2,3} & \sin\theta_{2,3} & \\ & -\sin\theta_{2,3} & \cos\theta_{2,3} & \\ & & & 1 \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & 1/\sqrt{3} & 2/\sqrt{6} & \\ & -2/\sqrt{6} & 1/\sqrt{3} & \\ & & & 1 \end{pmatrix}$$

$$\begin{aligned} \mathbf{A}^{(5)} &= \mathbf{G}_{2,3}^{(2)}\mathbf{A}^{(4)} = \begin{pmatrix} 1 & & & \\ & 1/\sqrt{3} & 2/\sqrt{6} & \\ & -2/\sqrt{6} & 1/\sqrt{3} & \\ & & & 1 \end{pmatrix} \begin{pmatrix} 2 & 3 & 2 \\ 0 & 5/\sqrt{3} & -2\sqrt{3} \\ 0 & 10/\sqrt{6} & 0 \\ 0 & 0 & 2\sqrt{2} \end{pmatrix} \\ &= \begin{pmatrix} 2 & 3 & 2 \\ 0 & 5 & -2 \\ 0 & 0 & 2\sqrt{2} \\ 0 & 0 & 2\sqrt{2} \end{pmatrix} \end{aligned}$$

Ref: An example of QR Decomposition, Che-Rung Lee, November

QR Decomposition Given's rotation example

6. Use a_{33} to eliminate a_{43}. $r_{3,4} = \sqrt{(2\sqrt{2})^2 + (2\sqrt{2})^2} = 4$. $\begin{cases} \cos\theta_{3,4} = a_{33}/r = 1/\sqrt{2}, \\ \sin\theta_{3,4} = a_{43}/r = 1/\sqrt{2}. \end{cases}$

$$\mathbf{G}_{3,4}^{(3)} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \cos\theta_{3,4} & \sin\theta_{3,4} \\ & & -\sin\theta_{3,4} & \cos\theta_{3,4} \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1/\sqrt{2} & 1/\sqrt{2} \\ & & -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

$$\begin{aligned} \mathbf{A}^{(6)} &= \mathbf{G}_{3,4}^{(3)}\mathbf{A}^{(5)} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1/\sqrt{2} & 1/\sqrt{2} \\ & & -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 2 & 3 & 2 \\ 0 & 5 & -2 \\ 0 & 0 & 2\sqrt{2} \\ 0 & 0 & 2\sqrt{2} \end{pmatrix} \\ &= \begin{pmatrix} 2 & 3 & 2 \\ 0 & 5 & -2 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Ref: An example of QR Decomposition, Che-Rung Lee, November

QR Decomposition Given's rotation example

7. The R-factor is $\begin{pmatrix} 2 & 3 & 2 \\ 0 & 5 & -2 \\ 0 & 0 & 4 \end{pmatrix}$; the Q-factor is

$$\begin{aligned} \mathbf{Q} &= \left(\mathbf{G}_{3,4}^{(3)}\mathbf{G}_{2,3}^{(2)}\mathbf{G}_{3,4}^{(2)}\mathbf{G}_{1,2}^{(1)}\mathbf{G}_{2,3}^{(1)}\mathbf{G}_{3,4}^{(1)}\right)^{-1} \\ &= \mathbf{G}_{3,4}^{(1)T}\mathbf{G}_{2,3}^{(1)T}\mathbf{G}_{1,2}^{(1)T}\mathbf{G}_{3,4}^{(2)T}\mathbf{G}_{2,3}^{(2)T}\mathbf{G}_{3,4}^{(3)T} \end{aligned}$$

$$= \begin{pmatrix} 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{pmatrix} = \mathbf{Q}$$

Ref: An example of QR Decomposition, Che-Rung Lee, November

Putting it together - SVD

Starting from the beginning with a matrix M, we want to derive - UWV^T

Using Householder transformations : $M = PBS$ [Step 1]

Using QR Decompositions : $B = Q_0^T Q_1^T \ldots Q_w^T W Q_w \ldots Q_1 Q_0$ [Step 2]

Substituting step 2 into 1 : $M = PQ_0^T Q_1^T \ldots Q_w^T W Q_w \ldots Q_1 Q_0 S$

With U being derived from : $U = PQ_0^T Q_1^T \ldots Q_w^T$

And V^T being derived from : $V^T = Q_w \ldots Q_1 Q_0 S$

Which results in the final SVD : $M = UWV^T$

SVD Applications

Calculation of the (pseudo) inverse :

[1] : Given $M = UWV^T$

[2] : Multiply by M^{-1} $M^{-1}M = M^{-1}UWV^T \rightarrow 1 = M^{-1}UWV^T$

[3] : Multiply by V $V = M^{-1}UWV^TV \rightarrow V = M^{-1}UW$

[4]* : Multiply by W^{-1} $VW^{-1} = M^{-1}UWW^{-1} \rightarrow VW^{-1} = M^{-1}U$

[5] : Multiply by U^T $VW^{-1}U^T = M^{-1}UU^T \rightarrow VW^{-1}U^T = M^{-1}$

[6] : Rearranging $M^{-1} = VW^{-1}U^T$

*Note – Inverse of a diagonal matrix is $\mathrm{diag}(a_1,\ldots,a_n)^{-1} = \mathrm{diag}(1/a_1,$

SVD Applications con't

Solving a set of homogenous linear equations i.e. *Mx = b*

Case 1 : b = 0

x is known as the nullspace of *M* which is defined as the set of all vectors that satisfy the equation *Mx = 0.* This is any column in V^T associated with a singular value (in *W*) equal to 0.

Case 2 : b != 0

Then we have : $Mx = b$

Which can be re-written as : $M^{-1}Mx = M^{-1}b \rightarrow x = M^{-1}b$

From the previous slide we know : $M^{-1} = VW^{-1}U^T$

Hence : $x = VW^{-1}U^T b$ which is easily solvable

SVD Applications con’t

Rank, Range, and Null space

- The **rank** of matrix **A** can be calculated from SVD by the number of nonzero singular values.
- The **range** of matrix **A** is The left singular vectors of **U** corresponding to the non-zero singular values.
- The **null space** of matrix **A** is The right singular vectors of **V** corresponding to the zeroed singular values.

$$A = UWV^T$$

Range Rank Null Space

$$\begin{bmatrix} \boxed{\begin{matrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{matrix}} & \begin{matrix} 0 \\ 0 \\ -1 \\ 0 \end{matrix} \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{5} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \begin{matrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \sqrt{0.2} & 0 & 0 \\ 0 & 0 & 0 \\ -\sqrt{0.8} & 0 & 0 \end{matrix} & \boxed{\begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & \sqrt{0.8} \\ 1 & 0 \\ 0 & \sqrt{0.2} \end{matrix}} \end{bmatrix}$$

SVD Applications con't

Condition number

- SVD can tell How close a square matrix A is to be singular.
- The ratio of the largest singular value to the smallest singular value can tell us how close a matrix is to be singular:

$$A = U \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \sigma_k \end{bmatrix} V^{\mathsf{T}} \qquad c = \frac{\sigma_1}{\sigma_k}$$

- A is singular if c is infinite.
- A is ill-conditioned if c is too large (machine dependent).

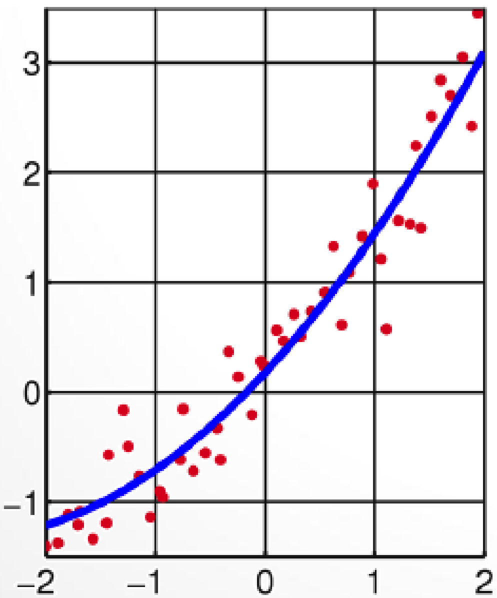
SVD Applications con't

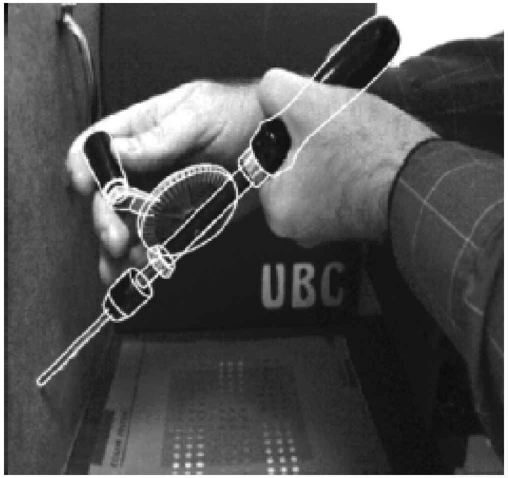
Data Fitting Problem

$$y = ax^2 + bx + c$$

$$\underbrace{\begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_N^2 & x_N & 1 \end{bmatrix}}_{S} \underbrace{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}_{\mathbf{a}} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix}}_{\mathbf{y}}$$

$$\mathbf{a} = V \cdot [\mathrm{diag}(1/\sigma_i)] \cdot (U^{\mathsf{T}}\ \mathbf{y})$$



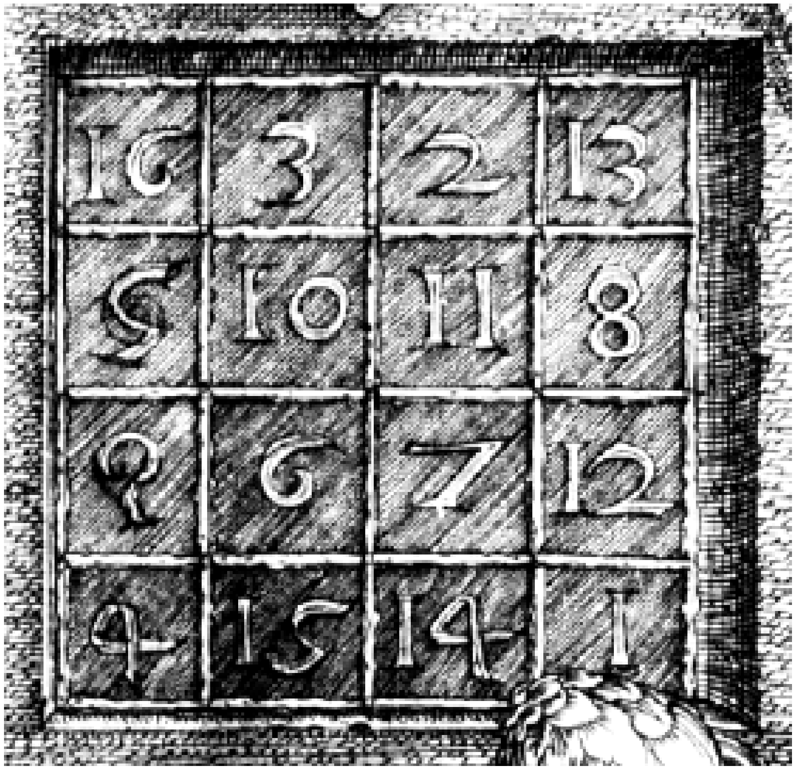


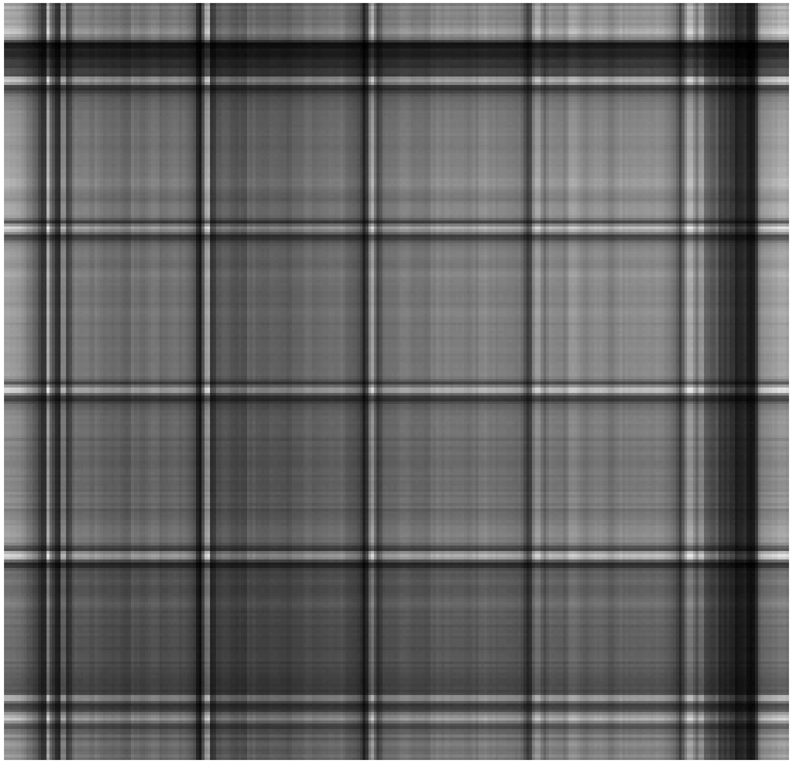
SVD Applications con't

Image processing

[U,W,V]=svd(A)

NewImg=U(:,1)*W(1,1)*V(:,1)'





SVD Applications con't

Digital Signal Processing (DSP)

- SVD is used as a method for noise reduction.

- Let a matrix *A* represent the noisy signal:
  - compute the SVD,
  - and then discard small singular values of *A*.

- It can be shown that the small singular values mainly represent the noise, and thus the rank-*k* matrix A_k represents a filtered signal with less noise.

Additional References

1. Golub & Van Loan – Matrix Computations; 3rd Edition, 1996
2. Golub & Kahan – Calculating the Singular Values and Pseudo-Inverse of a Matrix; SIAM Journal for Numerical Analysis; Vol. 2, #2; 1965
3. An Example of QR Decomposition, Che-Rung Lee, November 19, 2008