Application of the Singular Value Decomposition (SVD)

MA580

Presented by:
Ali Sadri
Secin Guncavdi
Min Yi

10-30-98

Singular Value Decomposition (SVD)

Singular Value Decomposition (SVD) of A if A is real *m-by-n* matrix such:

$$A=UDV^T \quad (1)$$

Where $U=[u_1,\ldots,u_m]$ m-by-m and $V=[v_1,\ldots,v_n]$ n-by-n are orthogonal and $D=[s_1,\ldots,s_p]$ where $p=min\{m,n\}$ and s_i are the singular values of A.

In this project we define five different problems that can be solved using the SVD algorithm. For each question a brief description of the problem is noted and a numerical result is presented.

Regularization of a Distorted Image by AWGN

Let assume H(t) is the operation to distort the image and N(t) is the Additive White Gaussian Noise (AWGN), the system can be expressed as a linear operation equation:

$$Y(t) = X(t) * H(t) + N(t)$$

Y(t) is the received or observed Image and X(t) is the original data. Let also assume that H(t) is a clear channel that is H(t) = I, the identity matrix. In this case our system is simplified as only an AWGN channel.

The Objection of this experiment is to reconstruct the original X(t) where the only information available is Y(t).

Solution

SVD can be applied to reconstruct the original signal. In this case our matrix Y(t) based on Eq. 1, Y(t) is matrix A. Therefore, we can compute the singular values for matrix Y(t). Let assume we have a 50-by-50-pixel image. Computing the Diagonal matrix $D=[s_1, \ldots,s_p]$ where $s_1>\ldots>s_p$ the smaller values of s_i corresponds to the noise subspace. Therefore, to eliminate the effect of the noise, we can replace the small s_i with zero and reconstruct the matrix $Y=UD\tilde{}V^T$ with the new $D\tilde{}$.

A simple experiment using a sphere was simulated and results show sufficient resolution enhancement and noise reduction.

Fig. 1 illustrates the original transmitted signal without the additive noise.

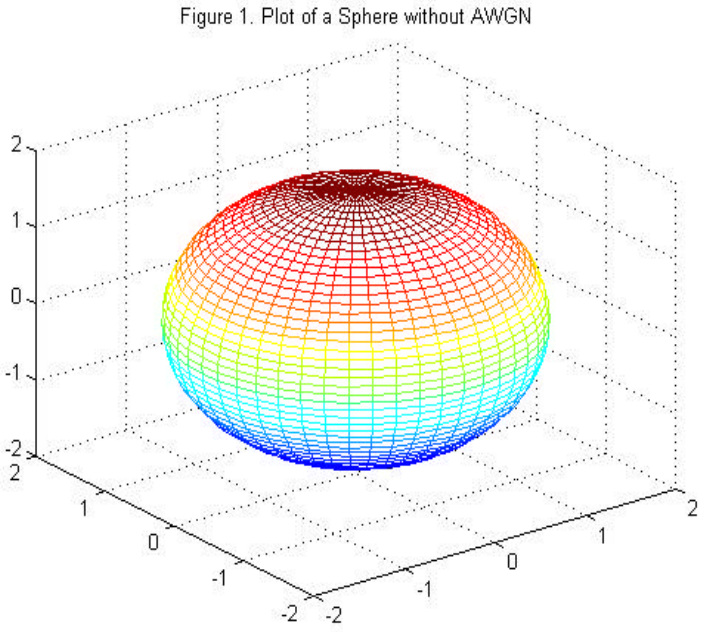
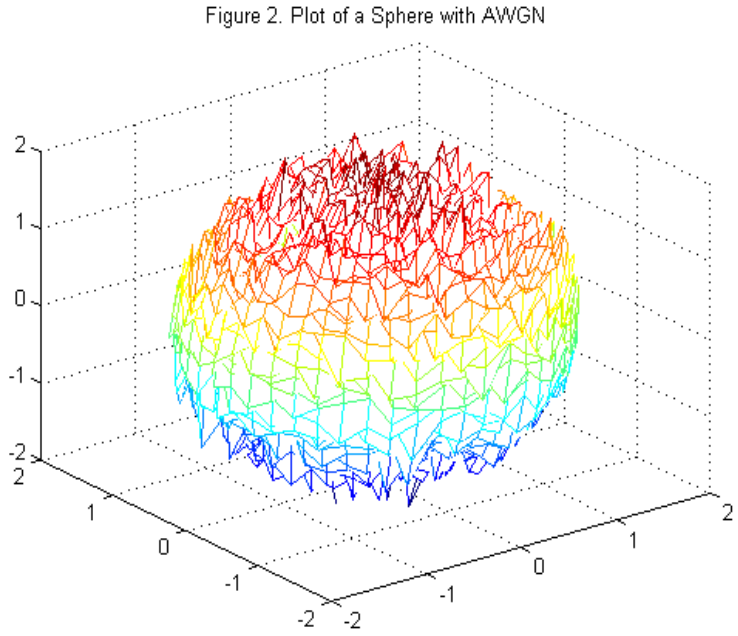
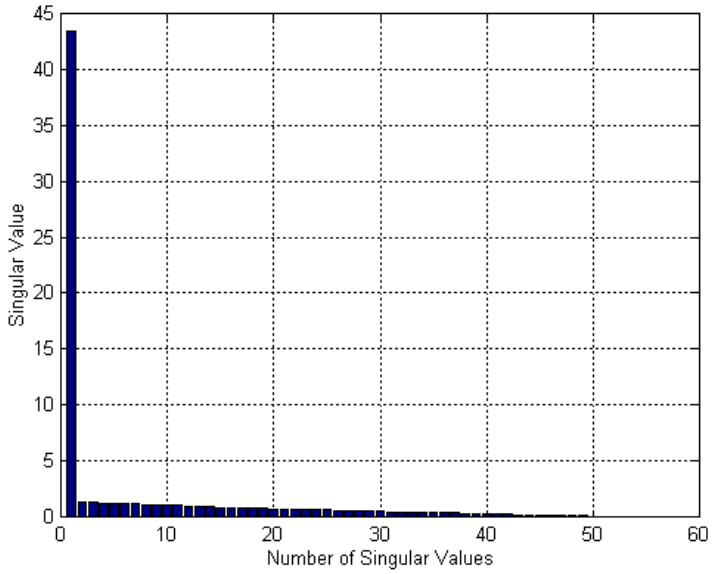


Fig 2, illustrates the sphere with AWGN.



To eliminate the noise, we have determined the Singular Values and compare it to a threshold. Fig. 3 illustrates the Singular Values of the sphere.

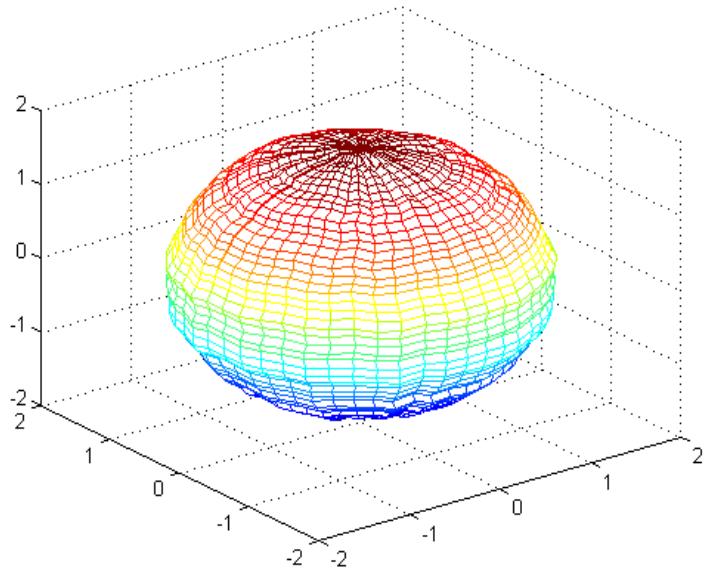
Figure 3. Plot of the Singular Values



It is clear from the above figure that the only suitable Singular Value is the first element of the diagonal matrix $s_1 >> \ldots > s_p$.

Finally the reconstructed matrix Y(t) after removing the AWGN is illustrated in Fig. 4.

Figure 4. Plot of a Sphere after Noise Reduction



By comparing Fig. 1 2, and 4, we can clearly notice the effect of AWGN in Fig. 2 and removal of noise in Fig. 4.

Control Application of Singular Value Decomposition

Linear feedback systems usually suffer from the inevitable disturbance from outside. Designing the forward and backward transfer functions is very important for robustness. The problem is choosing the right transfer functions and finding out how sensitive the system is.
Singular value decomposition (SVD) has become a ubiquitous element in modern control theory. The technical literature of the last ten years relies heavily on SVD for the treatment of linear multivariate control [1].
A typical application of SVD involves a multi-input multi-output feedback loop such as Fig. 5. The matrix **G** defines a linear transformation between two dimensioned spaces, to relate the multivariate signals **e** and **y**. **F** is a mapping in the reverse direction, from the space of **y** to the space of **e**. The signals **r**, **e**, **d**, **y** are usually modeled as vectors in $\Re^n$. System properties such as command following, noise immunity, and robustness are then analyzed in terms of the singular values of the loop transfer function, **GF**(s). Analogies to single-input-single-output system properties can be made in which the minimum or maximum singular values of **GF**(s) measure its size in a sense that generalizes the case of a scalar transfer function. These measures are more general since they allow that the amount of amplification can depend on the direction of the input signal.

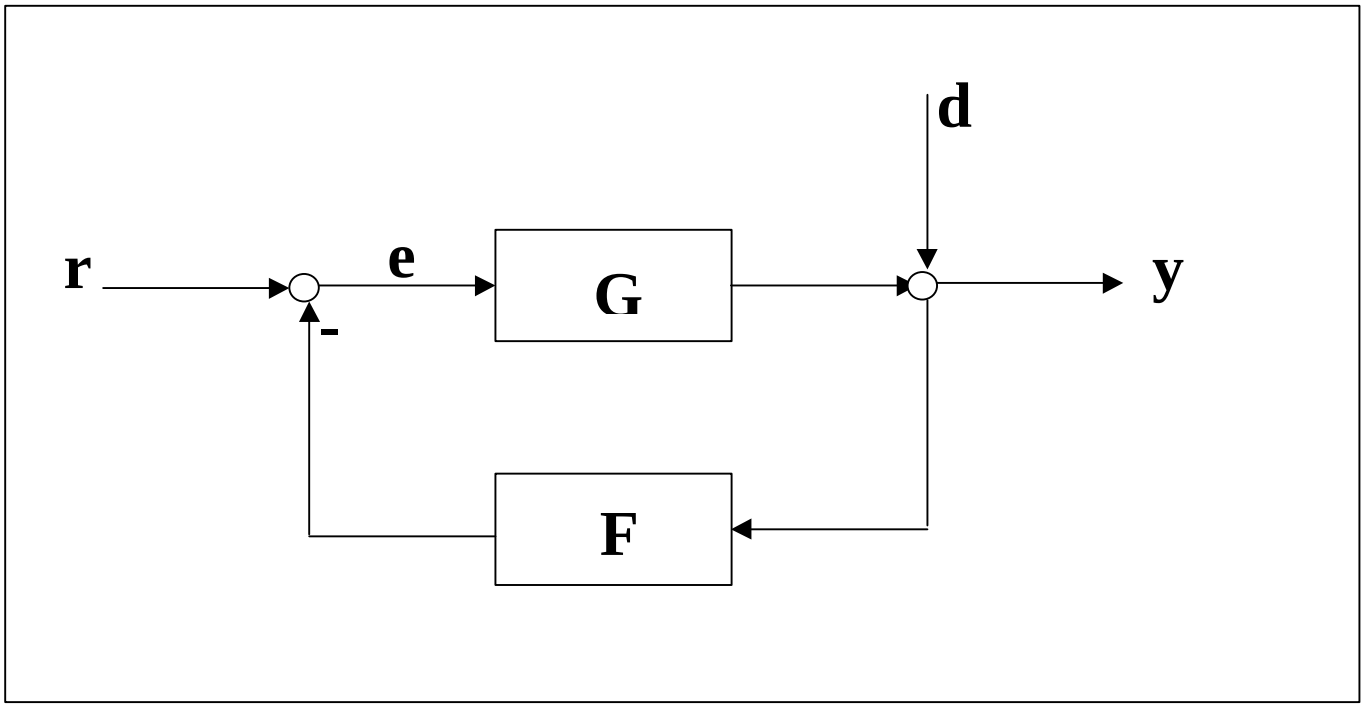


Figure 5. A Multidimensional feedback system

Consider loop shaping the feedback system of Fig. 1 to minimize the effects of the output disturbance, **d**.

Solving for **y**. we have:

e=r-Fy
y=d+Ge=d+G(r-Fy)
=d+Gr-GFy
(I+GF)y=d+Gr $\Rightarrow$ **$y=(I+GF)^{-1}d+(I+GF)^{-1}Gr$**

For the output to be insensitive to the disturbance, the sensitivity transfer function **$(I+GF)^{-1}$**, must be small in every direction. This requires in general that

$$||(I+GF)^{-1}||_2=\sigma_{max}((I+GF)^{-1})<<1$$

where σ_{max} is the maximum singular value of the matrix **$(I+GF)^{-1}$.**

Example:
r:100 samples from a full period sin. Wave of magnitude 100.

$$G=m*\begin{bmatrix} 1 & -0.5 & 0 & .. & 0 \\ 0 & 1 & -0.5 & .. & 0 \\ 0 & 0 & .. & : & : \\ : & 0 & .. & 1 & -0.5 \\ 0 & 0 & .. & 0 & 1 \end{bmatrix}_{100x100} \qquad d=\begin{bmatrix} 2 \\ 2 \\ 2 \\ : \\ 2 \end{bmatrix}_{100\,x1}$$

$$F=\begin{bmatrix} 0.5 & 0.2 & -0.1 & 0 & .. & 0 \\ 0 & 0.5 & 0.2 & -0.1 & .. & 0 \\ 0 & 0 & .. & .. & .. & : \\ 0 & 0 & .. & 0.5 & 0.2 & -0.1 \\ : & : & .. & 0 & 0.5 & 0.2 \\ 0 & 0 & .. & 0 & 0 & 0.5 \end{bmatrix}_{100x100}$$

Disturbance is kept constant and the multiplier of m is used to change the singular values of **$(I+GF)^{-1}$.**

The sensitivity criteria used the difference of the output signals with no disturbance and disturbance **d**. So;

y0=**=$(I+GF)^{-1}$Gr**
y1==**$(I+G'F)^{-1}$d+$(I+G'F)^{-1}$G'r** where **G'**=m***G;**

Sensitivity =var(**y1-y0**)/var(**y0**)
$Sigma_{max}=\sigma_{max}$ **$((I+G^{'}F)^{-1})$**

As a result sensitivity and σ_{max} are both functions of m. The calculations are done in MATLAB for 100 m values, which results in 100 sensitivity and σ_{max} values. Fig. 6, presents the results.

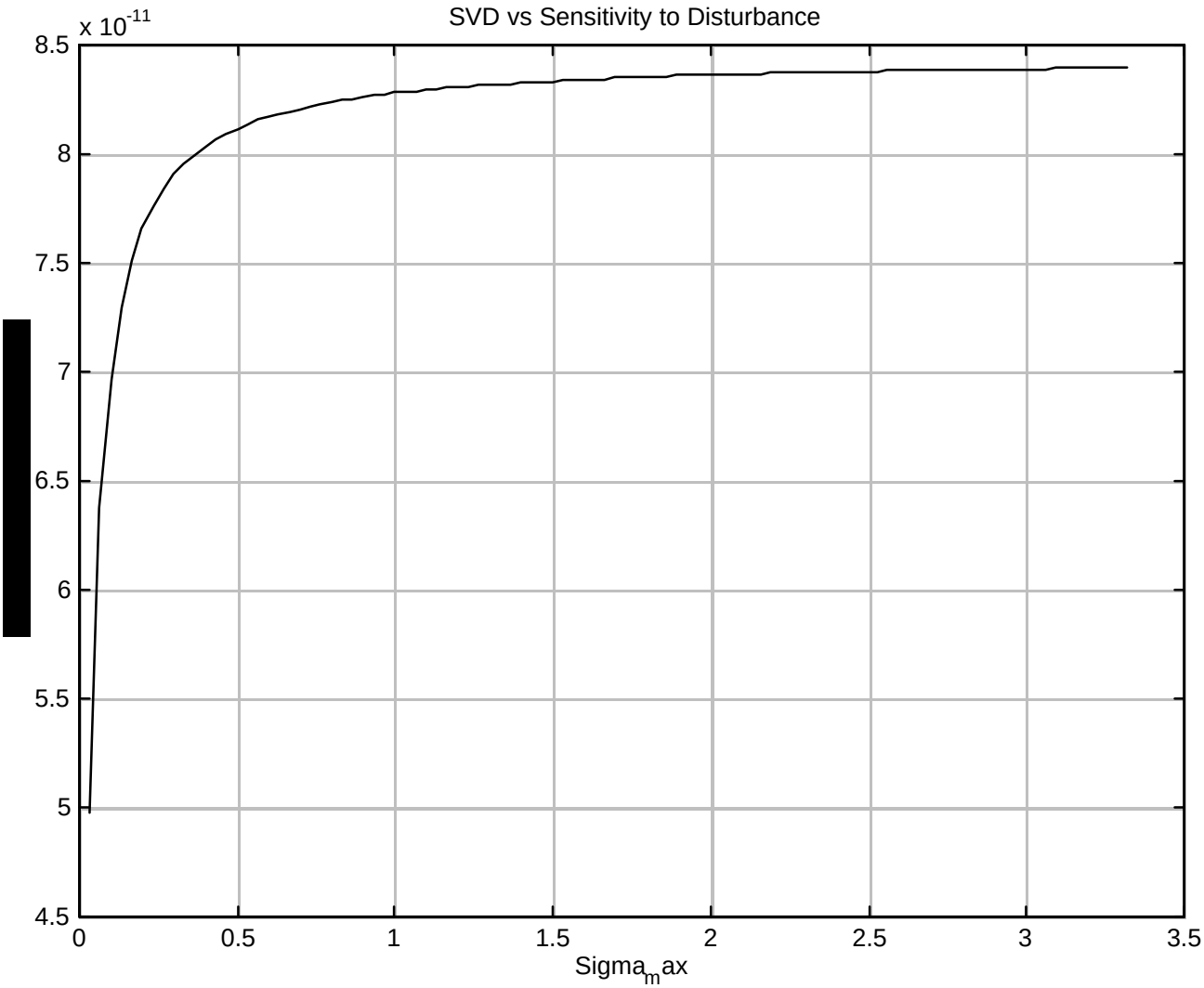


Figure 6. SVD Vs Sensitivity

As seen on the graph, the sensitivity of the output is highly dependent on the max. singular value of **$(I+G^{'}F)^{-1}$.** For $sigma_{max}<<1$, sensitivity is low compared to $sigma_{max}>>1$. The drastic change occurs while $sigma_{max}$ is increasing when it is less than 1. As a result, using SVD for the feedback control systems, the stability and sensitivity behavior of the system can be controlled.

[1]:Hart, George W., Multidimensional Analysis, Springer-Verlag New York 1995.

The conditioning of the reference node choice in an electrical network

The analysis of circuits, in general, requires writing and solving a set of simultaneous equations. These equations are mathematical expressions of Kirchhoff's laws and voltage current relationships applied to the circuit and its components. Systematic procedures are necessary to select the proper set of independent variables and write the equations of a given circuit. Assume the N unknowns are the terminal voltages of the circuit. Circuit equations are written by choosing one of the terminals as the reference node, since current thru an resistor is defined by the difference between two ends divided by the resistance value (Ohm's law).
The solution to the circuit is obtained by solving these equations, however the problem arises when numerical solutions are applied, i.e-finite precision is used. The existence and algebraic properties of singular matrix representations for N-terminal circuits have been described during nearly 60 years. It will be shown how from a practical point of view certain network descriptions are better suited when the accuracy of the computations is pushed to extreme value [2].
The resulting set of equations represents a linear system as:

$$\mathbf{Y^*[v_1^*\ldots v_t^*\ldots v_N^*]^t=[i_1^*\ldots i_s^*\ldots i_N^*]^t \text{ or } Y^*.v^*=i^*}$$
$$\mathbf{Y[v_1\ldots v_t\ldots v_N]^t=[i_1\ldots i_s\ldots i_N]^t \text{ or } Y.v=I}$$

The first equation is the indefinite representation, in which all-terminal potentials $\mathbf{v_k^*}$ **and currents** $\mathbf{i_k^*}$ **occur.** The second equation is the definite representation for which node voltage v_t has been chosen as a reference voltage and Kirchoff Current Law equation at node s has been dropped. Usually s=t. Y is obtained by simply deleting row s and column t in Y^*. These trivial algebraic relationships can be invoked to qualify both descriptions are equivalent. However, it is demonstrated that numerically, this statement is incorrect. It turns out that, even for surprisingly simple circuits, one of the matrices Y or Y^* may be ill conditioned, whereas the other is well conditioned. This is where SVD is used, since:

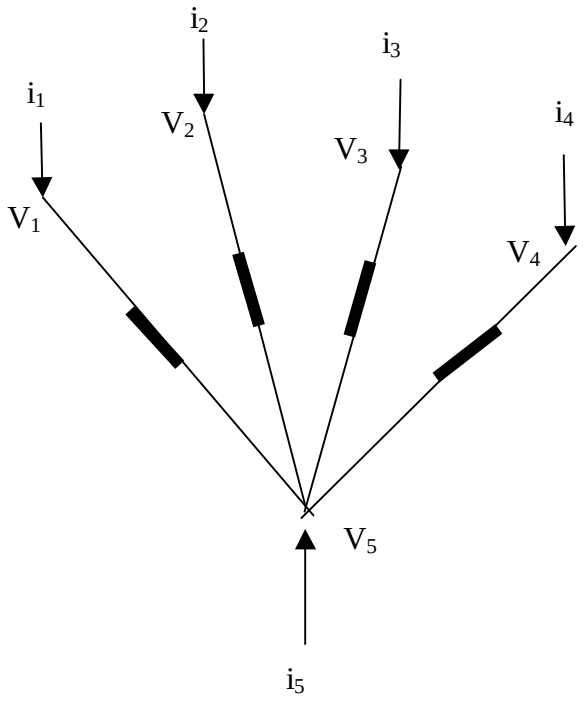
$$\mathbf{K(A)=\sigma_{max}/\sigma_{min}}$$

Where σ is a singular value of the matrix A, K(A) is the condition number of the matrix A.

So, comparing condition numbers of different Y matrices obtained by choosing different reference node and Y^*, one can choose the one with the smallest conditioning number hence getting the best conditioned matrix to solve.

Example:

Consider the following circuit:



Where each ▬ represents a resistance of value R.

Using the Kirchoff's law at nodes 1, 2, 3, 4 and 5, we get:

1- $1/R(V_1-V_5)=i_1$

2- $1/R(V_2-V_5)=i_2$

3- $1/R(V_3-V_5)=i_2$

4- $1/R(V_4-V_5)=i_3$

5- $1/R(V_5-V_1)+(V_5-V_2)+(V_5-V_3)+(V_5-V_4)=i_5$

Which corresponds to the system;

$$1/R\begin{bmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 \\ -1 & -1 & -1 & -1 & 4 \end{bmatrix}\begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \end{bmatrix} = \begin{bmatrix} i_1 \\ i_2 \\ i_3 \\ i_4 \\ i_5 \end{bmatrix}$$

Or, $Y^*.V=I$, Choose R=1;

Lets choose the reference node as k=2, so that Y will be obtained by deleting row 2 and column 2.

$$Y = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ -1 & -1 & -1 & 4 \end{bmatrix}$$

$$SVD(Y) = \begin{bmatrix} 4.7913 \\ 1.000 \\ 1.000 \\ 0.2087 \end{bmatrix}$$

$K(A)=\sigma_{max}/\sigma_{min}=22.9564$

Lets choose the reference node as k=5, so that Y will be obtained by deleting row 5 and column 5.

$$Y = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Clearly all singular values of Y is 1. So K(Y)=1.

So choosing reference node has an important effect on the condition number of the matrix. With the increasing number of nodes, these become more important. Every element in the circuit has a tolerance, which is a noise for the system and finite precision is used for the numerical analysis of the circuits. So ill-conditioned matrices i.e large condition numbers should be avoided. Using SVD, condition number can be found using the formula above, which allows us to decide which node to choose as a reference to get a well-conditioned matrix.

[2] Deprettere Ed. F, SVD and Signal Processing, North-Holland, Netherlands, 1988.

Computing Low-Rank Approximations

Given an $m \times n$ matrix A of rank r ($r \le \min(m,n)$), sometimes we need to get a matrix A_k of a lower rank k (1<=k<=r) from A. How can we make A_k the best approximation of the matrix A by matrices of lower rank? The problem goes to how to get the minimum of the 2-norm or the Frobenius norm of (A-B), where B is any matrix of rank k, i.e., $\min_{rank(B)=k} \|A-B\|_{2 or Frobenius-norm}$. The SVD can be used as a tool to solve this problem.

If we have got the singular values of A which are $\sigma_1 \ge ... \ge \sigma_r > \sigma_{r+1} = ... = 0$ and also the right singular vectors $u_1 \dots u_n$ of A and the left singular vectors $v_1 \dots v_m$ of A.

Then $A_k = \sum_{i=1}^{k} \sigma_i v_i u_i^T$

A_k Satisfies $A_k = \min_{rank(B)=k} \|A-B\|_{2 or Frobenius-norm}$ and it is exactly what we need.

Now let's figure out why SVD is working in this way.

$A_k = \sum_{i=1}^{k} \sigma_i v_i u_i^T = V \,\mathrm{diag}(\sigma_1,...,\sigma_k,0,...,0)\, U^T$,

So $V^T A_k U = \mathrm{diag}(\sigma_1,...,\sigma_k,0,...,0) \Rightarrow \mathrm{rank}(A_k) = k$

Thus, $V^T (A - A_k) U = \mathrm{diag}(0,...,0,\sigma_{k+1},\sigma_{k+2},...) \Rightarrow \|A - A_k\|_2 = \sigma_{k+1}$

Suppose there is some B with rank (b)=k$\le r$. Let $\{x_1, x_2, \dots, x_{n-k}\}$ are the basis of the null space of B. The dimension of the null space of B, i.e., N(B), is n-k and the dimension of the space span by $\{u_1, u_2, \dots, u_{k+1}\}$ is k+1. Both of them are the subspaces of R^n.

So N(B)$\bigcap$ span$\{u_1, u_2, \ldots, u_{k+1}\} \neq \{0\}$.

Let z be a unit 2-norm vector in N(B)$\bigcap$span $\{u_1, u_2, \ldots, u_{k+1}\}$. Since z is in N(B), so Bz=0.

Also, $Az=(\sum_{i=1}^{r} \sigma_i v_i u_i^T)z=\sum_{i=1}^{r} \sigma_i (u_i^T z)v_i \in R^m$

So $\|A-B\|_2^2 \geq \|(A-B)z\|_2^2 = \|Az\|_2^2 = \sum_{i=1}^{r} \sigma_i^2 (u_i^T z)^2 \geq \sigma_{k+1}^2 \sum_{i=1}^{k+1} (u_i^T z)^2 = \sigma_{k+1}^2$

Then we get the conclusion that A_k is just the minimum of the 2-norm or the Frobenius norm of (A-B), where B is any matrix of rank k. SVD method works well on this problem.

Numerical example:

$$A = \begin{bmatrix} 3 & -3 & 3 \\ 1 & 5 & -7 \\ -2 & 2 & -2 \\ 10 & -12 & 5 \end{bmatrix}$$

After computing the singular value decomposition, we get matrix U with the right singular vectors as its columns vectors, matrix V with the left singular vectors as its column vectors and

$$U = \begin{bmatrix} 5.2655e-01 & -6.7723e-01 & 5.1391e-01 \\ -7.3188e-01 & -5.3546e-02 & 6.7932e-01 \\ 4.3254e-01 & 7.3382e-01 & 5.2385e-01 \end{bmatrix}$$

matrix S with singular values on its diagonal.

$$V = \begin{bmatrix} 2.7596e-01 & 5.0273e-02 & 7.8334e-01 & -5.5470e-01 \\ 3.3514e-01 & -9.2531e-01 & 1.7745e-01 & -7.6328e-17 \\ -1.8398e-01 & -3.3515e-02 & -5.2223e-01 & -8.3205e-01 \\ 8.8186e-01 & -3.7437e-01 & -2.8665e-01 & 1.6653e-16 \end{bmatrix}$$

$$S = \begin{bmatrix} 1.832e+01 & 0 & 0 \\ 0 & 6.5726e+00 & 0 \\ 0 & 0 & 1.3727e+00 \\ 0 & 0 & 0 \end{bmatrix}$$

Now we need to get a matrix A_2 with rank 2 which holds ‘as much of the energy’ of A as possible.

In the way explained above, A_2 is as following:

$$A_2 = \begin{bmatrix} 2.4474e+00 & -3.7305e+00 & 2.4367e+00 \\ 8.7482e-01 & 4.8345e+00 & -7.1276e+00 \\ -1.6316e+00 & 2.4870e+00 & -1.6245e+00 \\ 1.0202e+01 & -1.1733e+01 & 5.2061e+00 \end{bmatrix}$$

$\|A - A_2\|_2 = 1.3727e+00 = s_3$

Linear Least Squares Problem

The least square data fitting can be described as the solution of an over determined system of equations Ax=b. In this case, the system of equation has more equations than unknowns, i.e., more rows than columns. Suppose the system has m equations and n unknowns with m>n. We hope to find a vector x which satisfies Ax=b. Because range (A) is at most has dimension n and b is an m -dimensional, so b doesn't exist in the range of A. Such a problem doesn't have a solution. So we only can solve the equation system by minimizing the 2-norm of Ax-b (the residual), i.e., $\min_{x \in R^n} \|Ax - b\|_2^2$. The residual can be made very small by choosing suitable x, but it can not be made zero. We can use SVD to solve the above problem.

Let A= $V\,\mathrm{diag}.(\sigma_1,...,\sigma_r,0,...,0)\,U^T$

Consider $\|Ax - b\|_2^2 = \|V^T(Ax - b)\|_2^2 = \|V^T AUU^T x - V^T b\|_2^2$

(Because $\|V^T(Ax - b)\|_2$ =the square root of (the largest eigenvalue of

$(V^T(Ax - b))^T V^T(Ax - b) = (Ax - b)^T VV^T(Ax - b) = (Ax - b)^T(Ax - b) = \|Ax - b\|_2$

Let z= $U^T x$ and c= $V^T b$. We write z and c as partitions, z= $\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \in R^r$, c= $\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \in R^r$.

So $\|Ax - b\|_2^2 = \|\mathrm{diag}(s_1 \quad ... \quad s_r \quad 0 \quad ... \quad 0)z - c\|_2^2 = \left\| \begin{bmatrix} \sum z_1 - c_1 \\ -c_2 \end{bmatrix} \right\|_2^2 = \|\sum z_1 - c_1\|_2^2 + \|c_2\|_2^2$

If $z_1 = \sum^{-1} c_1$, then

$\|\sum z_1 - c_1\| = 0$, so $\|Ax - b\|_2^2$ is minimized.

We just choose $z_o = \begin{bmatrix} \sum^{-1} c_1 \\ 0 \end{bmatrix}$, then $x_o = Uz_o = U\begin{bmatrix} \sum^{-1} c_1 \\ 0 \end{bmatrix} = U\begin{bmatrix} \sum^{-1},0 \\ 0,0 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = U\begin{bmatrix} \sum^{-1},0 \\ 0,0 \end{bmatrix} V^T b$.

If there is any other least square solution x^{∞} , then z^{∞} must have the form $\begin{bmatrix} \sum^{-1} c_1 \\ z_2 \end{bmatrix}$, so $\left\|x^{\infty}\right\|_2^2 = \left\|Uz^{\infty}\right\|_2^2 = \left\|\sum^{-1} c_1\right\|_2^2 + \left\|z_2\right\|_2^2 \geq \left\|x_o\right\|_2^2$.

Numerical Example

Let matrices "A" and "b" to be:

$$A = \begin{bmatrix} 3 & 1 & 2 \\ 4 & 12 & -3 \\ -2 & 15 & -17 \\ 1 & -4 & 7 \end{bmatrix} \qquad b = \begin{bmatrix} 2 \\ -9 \\ 14 \\ 11 \end{bmatrix}$$

After computing the singular value decomposition, we get matrix U with the right singular vectors as its columns vectors, matrix V with the left singular vectors as its columns vectors and matrix S with singular values on its diagonal.

$$S = \begin{bmatrix} 2.6195e+01 & 0 & 0 \\ 0 & 88878e+00 & 0 \\ 0 & 0 & 1.3473e+00 \\ 0 & 0 & 0 \end{bmatrix}$$

$$V = \begin{bmatrix} 2.6882e-02 & -3.9459e-01 & 8.1474e-01 & 4.2401e-01 \\ -4.0841e-01 & -8.3022e-01 & -2.3266e-01 & -2.9967e-01 \\ -8.6329e-01 & 2.8900e-01 & -4.5575e-02 & 4.1126e-01 \\ 2.9531e-01 & -2.6741e-01 & -5.2915e-01 & 7.4919e-01 \end{bmatrix}$$

$$U = \begin{bmatrix} 1.7901e-02 & -6.0196e-01 & 7.9833e-01 \\ -7.2550e-01 & -5.5724e-01 & -4.0390e-01 \\ 6.8799e-01 & -5.7195e-01 & -4.4669e-01 \end{bmatrix}$$

S1 is the pseudo-inverse of S

$$S_1 \begin{bmatrix} 3.8175e-01 & 0 & 0 & 0 \\ 0 & 1.1251e-01 & 0 & 0 \\ 0 & 0 & 7.4223e-01 & 0 \\ & & & \end{bmatrix}$$

We get

$$X = \begin{bmatrix} -2.157e+00 \\ 4.7325e-01 \\ 2.7160e-01 \end{bmatrix}$$

And

$\|Ax - b\|_2 =$ **1.7544e+01**