Singular Value Decomposition

- General case of $[K]\{u\} = b$ when $[K]$ $m \times n$, $m \geq n$

  can we define $[K]^{-1}$ such that $\{u\} = [K]^{-1}\{b\}$?

- Any $[K]$ can be factored as

$$[K] = [U][\mathrm{diag}(w)][V^T]$$

such that

- $[K]$ and $[U]$ are $m \times n$
- $[\mathrm{diag}(w)]$ and $[V]$ are $n \times n$
- Columns of $[U]$ are orthogonal, $U^T U = I$
- Columns of $[V]$ are orthogonal, $V^T V = I$

  also $V V^T = I$ since square
- IF $m = n$, $U U^T = I$
- All $w_i \geq 0$

e.g.

$$[K] = \begin{bmatrix} | & | & & | \\ U_1 & U_2 & \cdots & U_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} w_1 & & & 0 \\ & w_2 & & \\ & & \ddots & \\ 0 & & & w_n \end{bmatrix} \begin{bmatrix} \cdots & V_1 & \cdots \\ \cdots & V_2 & \cdots \\ & \vdots & \\ \cdots & V_n & \cdots \end{bmatrix}$$

$\{U\}_i$ = Left singular vectors of $[K]$

$\{V\}_i$ = Right singular vectors of $[K]$

w_i = Singular values (typically ordered largest to smallest such that $K = w_1/w_n$)

$$[K][V] = [U][\mathrm{diag}(w)] \quad \text{since } V^TV = I$$

or for each i

$$[K]\{V\}_i = w_i\{U_i\}$$

i.e. $[K]$ maps $\{V\}_i$ onto $\{U_i\}$ with scaling w_i

A. <u>Square, nonsingular case</u> $(m=n)$

$$[K]^{-1} = [V][\mathrm{diag}(1/w)][U]^T$$

so $$\{u\} = [V][\mathrm{diag}(1/w)[U]^T\{b\}$$

- Columns of $[V]$ constitute the natural basis for $\{u\}$
  i.e. $\{u\} = [V]\{c\} \Leftrightarrow \{c\} = [V]^T\{u\}$
- Columns of $[U]$ are natural basis for $\{b\}$
  i.e. $\{b\} = [U]\{d\} \Leftrightarrow \{d\} = [U]^T\{b\}$
- Inversion projects $\{b\}$ onto U-space, multiplies by $(1/w)$ and reassembles $\{u\}$ in V-space

$$\{u\} = \sum_{i=1}^{N} \{V\}_i \frac{\{U\}_i \cdot \{b\}}{\omega_i} \iff c_i = \frac{d_i}{\omega_i}$$

when $[K]$ symmetric $[U] = [V]$, $\omega \to \lambda$

- Relative to noise in $\{b\}$

$$[Cov(u)] = [K^{-1}][Cov(b)][K^{-T}]$$

$$= [V][diag(1/\omega)][U]^T[Cov(b)][U][diag(1/\omega)][V]^T$$

$$[Cov(c)] = [diag(1/\omega)][Cov(d)][diag(1/\omega)]$$

IF $[Cov(b)] = \sigma^2 I = [Cov(d)]$

then we have

$$[Cov(u)] = \sigma^2 [V][diag(1/\omega)^2][V]^T$$

$$[Cov(c)] = \sigma^2 [diag(1/\omega)^2]$$

- Here $\{b\}$ is projected onto $[U]$ (rather than V)
- Noise present in $\{U_i\}$ (from $\{b\}$) will show up in in $\{u\}$ through $\{V\}_i$ amplified by $1/\omega_i$
- Small ω_i are noise amplifiers relative to lage ones making $\underbrace{K}_{\text{Condition \#}}$ the measure of noise distortion

B. Square Singular Case

IF a singular value is zero, i.e. $w_N = 0$

- If $\{b\}$ orthogonal to $\{U\}_N$, i.e. $\{b\} \cdot \{U\}_N = 0$ (or $d_N = 0$)

then $$\{u\} = \sum_{i=1}^{N-1} \{V\}_i \frac{\{U\}_i \cdot \{b\}}{w_i}$$

Same as non-singular case, except $\{V\}_N$ left out (Not needed to satisfy $[K]\{u\} = \{b\}$)

<u>But</u> solution <u>not unique</u> ... can add any amount of $\{V\}_N$ to solution without changing right hand side since by defn

$$[K]\{V\}_i = w_i \{U\}_i$$

and

$$[K]\{V\}_N = 0$$

∴ can have family of solutions

$$\{u\} = \sum_{i=1}^{N-1} \{V\}_i \frac{\{U_i\} \cdot \{b\}}{w_i} + \alpha \{V\}_N$$

$$= \sum_{i=1}^{N-1} \{V\}_i \frac{d_i}{w_i} + \alpha \{V\}_N$$

all of which satisfy $[K]\{u\} = \{b\}$

Note ... common to set $\alpha = 0$ since this is the minimum variance solution

$$Var(u) = \sum_{i=1}^{N-1} \left(\frac{\{U\}_i \cdot \{b\}}{\omega_i} \right)^2 + \cancel{\alpha^2} \xrightarrow{} 0 \Big\}$$ operationally same as setting $\frac{1}{W_N} \to 0$

• But when $d_N \equiv \{U\}_N \cdot \{b\} \neq 0$ (more likely, especially due to R/o)

have same options ... either include an arbitrary amount of $\{V\}_N$ in sol'n or remove it

$$\{u\} = \sum_{i=1}^{N-1} c_i \{V\}_i + \alpha \{V\}_N$$

$$[K]\{u\} = \{b\} = \sum_{i=1}^{N} d_i \{U\}_i$$

$$[K] \left(\sum_{i=1}^{N-1} c_i \{V\}_i + \alpha \{V\}_N \right) - \sum_{i=1}^{N} d_i \{U\}_i \equiv \{r\}$$

$$\sum_{i=1}^{N-1} c_i \underbrace{[K]\{V\}_i}_{w_i \{U\}_i} + \alpha \underbrace{[K]\{V\}_N}_{= 0 \text{ since } W_N = 0} - \sum_{i=1}^{N} d_i \{U\}_i$$

$$\{r\} = \sum_{i=1}^{N-1} c_i w_i \{U\}_i - \sum_{i=1}^{N} d_i \{U\}_i$$

$$= \sum_{i=1}^{N-1} (c_i w_i - d_i) U_i - d_N \{U\}_N$$

$$Var(r) = \sum_{i=1}^{N-1} (c_i w_i - d_i)^2 + d_N^2$$

the minimum variance sol'n is $c_i = \frac{d_i}{w_i}$

doesn't matter what α is, i.e. $\alpha = 0$ adds nothing to the residual

$\therefore$ Conclude if $w_N = 0$, treat it as if infinite, i.e. eliminate Nth term and will get minimum variance sol'n, i.e. keep $\{V\}_N$ out of sol'n – it is undetermined and ignore the presence of $\{U\}_N$ in $\{b\}$, there is no way to reduce its presence in the residual

C. <u>Square, Nearly-Singular Case</u> (i.e. $w_N \approx 0$)

Options: include w_N in the computation <u>or</u> treat it as if zero

- If retain, then have full N dimensional basis for $\{b\}$ and $\{U\}$, but inversion becomes noise amplifier

- If remove, then avoid noise amplification, <u>but</u> sol'n can't have any $\{V\}_N$ content and any forcing in $\{U\}_N$ is ignored. If $\{U\}_N$ an important mode of forcing, then in trouble because noise in its specification will dominate

- In practice ... consider all $w_i = 0$ by defining a cutoff based on condition number such that w_i's below the cutoff are treated as zero

- Eliminates any $\{V\}_i$ in solution for these w_i and ignores any $\{U\}_i$ in the forcing (rather than amplifying it)

<u>Over-Determined Case</u> $m > n$ (more eqn's than unknowns)

- will always have nonzero residual $\{r\} = [K]\{u\} - \{b\}$ but SVD gives minimum variance residual

- columns of $[V]$ are complete, orthonormal basis for $\{u\}$ $\Rightarrow$ $\{u\} = [V]\{c\} \Leftrightarrow \{c\} = [V]^T\{u\}$

- columns of $[U]$ an incomplete basis for $\{b\}$

$$\{d\} = [U]^T\{b\} \Leftrightarrow [U]\{d\} = \{b\} - \{b'\}$$

where $\{b'\}$ lies outside $[U]$ space $\Rightarrow [U]^T\{b'\} = 0$

- Since $[K]\{V\}_i = w_i\{U_i\}$ and $\{u\}$ is completely contained in $[V]$, then $[K]\{u\}$ completely contained in $[U]$, $\therefore$ $\{b'\}$ cannot be reached by any $\{u\}$.

then $\{u\} = \sum_{i=1}^{N} c_i\{V\}_i$

$$\{r\} = \sum_{i=1}^{N} (c_i w_i - d_i)\{U_i\} - \{b'\}$$

$$Var(r) = \sum_{i=1}^{N} (c_i w_i - d_i)^2 + Var(b')$$

and minimum variance sol'n is $c_i = \frac{d_i}{w_i}$

$$\{u\} = \sum_i \{V\}_i \frac{\{U\}_i \cdot \{b\}}{w_i}$$

Under-Determined Case $m < n$

- add equations $0 \cdot u = 0$ to make $m = n$, which generates $n-m$ modes with $w_i = 0$
- handle as above, eliminate the $w_i = 0$

So in every case we have the same procedure

- eliminate all singular and nearly-singular modes ($w_i \approx 0$) from the calculations
- results in a reduced-rank system which prevents noise amplification by leaving it uninverted