Matrix Inversion and Inverse Noise

- So far w/ FD or FE have equivalence between

$$[K]\{u\}=\{b\} \iff \{u\}=[K^{-1}]\{b\}$$

  - Assume K^{-1} exists (don't actually compute it!)
  - K^{-1} generally "full" even if $[K]$ sparse
    i.e. each u_i is dependent on <u>all</u> $\{b\}$
  - $[K]$ is square and full rank (all rows linearly independent)
    → each row an FD molecule or galerkin eqn for ϕ_i
  - b includes BCs, forcing

- What if $\{b\}$ has uncertainty ... e.g. based on data w/ noise

then $\{b\}=\{\bar{b}\}+\{\tilde{b}\}$

"mean" or "true" value

unknown component due to measurement noise or other uncertainty

"System Noise"

and we expect

$$\{u\}=\{\bar{u}\}+\{\tilde{u}\}$$

$\bar{u} \equiv$ Mean response

$\tilde{u} \equiv$ Inverse Noise

since $[K]\{\bar{u}\}=\{\bar{b}\}$ <u>or</u> $\{\bar{u}\}=[K^{-1}]\{\bar{b}\}$

$[K]\{\tilde{u}\}=\{\tilde{b}\}$ <u>or</u> $\{\tilde{u}\}=[K^{-1}]\{\tilde{b}\}$

so $\bar{u}$ computed from $\bar{b}$ and $\tilde{u}$ from $\tilde{b}$

where $\{\overline{\tilde{u}}\} = 0$ since $\{\overline{\tilde{b}}\} = 0$ (Noise is zero mean)

- Statistical characteristics of $\tilde{u}$... how related to those of $\tilde{b}$ (i.e. system noise)

• <u>Covariance</u>

$$[Cov(\tilde{u})] \equiv \overline{\tilde{u}\tilde{u}^T}$$

(recall $\tilde{u}$ is $n \times 1$, $\tilde{u}^T$ is $1 \times n$, $\tilde{u}\tilde{u}^T$ is $n \times n$)

But $\tilde{u}\tilde{u}^T = [K^{-1}]\{\tilde{b}\}\{\tilde{b}^T\}[K^{-T}]$

so $\overline{\tilde{u}\tilde{u}^T} = [K^{-1}]\overline{\{\tilde{b}\}\{\tilde{b}^T\}}[K^{-T}]$

or

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

- IF noise is uncorrelated: $Cov(\tilde{b}) = \sigma^2 I$ ← noise variance

then $[Cov(\tilde{u})] = \sigma^2 \underbrace{[K^{-1}][K^{-T}]}$

acts like a noise filter

key: does it amplify or suppress?

Can partition $\tilde{b}$ as

$$\{\tilde{b}\} = \{\tilde{b}\}_D + \{\tilde{b}\}_N + \{\tilde{b}\}_I$$

due to Type I (Dirichlet) — due to Type II (Neumann) — due to interior forcing

$$\text{then } \{\tilde{u}\} = \underbrace{[K^{-1}]\{\tilde{b}\}_D}_{\{\tilde{u}\}_D} + \underbrace{[K^{-1}]\{\tilde{b}\}_N}_{\{\tilde{u}\}_N} + \underbrace{[K^{-1}]\{b\}_I}_{\{\tilde{u}\}_I}$$

3 sources of variability, IF independent (statistically)

$$[Cov(\tilde{b})] = [Cov(\tilde{b})]_D + [Cov(\tilde{b})]_N + [Cov(b)]_I \quad \underline{\text{additive}}$$

and inverse noise covariance is also <u>additive</u>

$$[Cov(\tilde{u})] = [K^{-1}][Cov(b)]_D[K^{-T}] + [K^{-1}][Cov(b)]_N[K^{-T}] + [K^{-1}][Cov(b)]_I[K^{-T}]$$

$$= [Cov(\tilde{u})]_D + [Cov(\tilde{u})]_N + [Cov(\tilde{u})]_I$$

IF sources of variability <u>not</u> independent, get additional cross-terms and separation into effects due to Type I, Type II, Interior not that useful.

- Variance

  - diagonals of covariance matrix = variances of individual elements of the vector involved

  - Sum of diagonals ≡ Trace, so $Trace[Cov(\tilde{u})]$ is scalar measure of variability of u

$$Var(\tilde{u}) = \overline{\{\tilde{u}\}^T\{\tilde{u}\}} = Tr[Cov(\tilde{u})] = \underbrace{\sum_{i=1}^{N} \overline{\tilde{u}_i^2}}_{\text{Divide by } N \text{ get MSS (Mean squared size)}}$$

$$= \overline{\{\tilde{b}\}^T[K^{-T}][K^{-1}]\{\tilde{b}\}}$$

- Noise Models

  - often we analytical models of input variability

e.g. $\overline{\tilde{b}_i\tilde{b}_j} = \sigma^2(1 + r_{ij}/\ell)\, e^{-r_{ij}/\ell}$

r_{ij} ≡ separation distance between locations (i.e. nodes) i and j

σ ≡ standard deviation

ℓ ≡ correlation length $1\ell \sim 75\%$; $2\ell \sim 40\%$; $4\ell \sim 10\%$

$[Cov(\tilde{b})]$ will be symmetric, full, strong diagonal terms

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

← full, diagonal terms = $\overline{\tilde{u}^2}$

<u>Example</u> : Node 2 has point source with strength 1 ± 0.5
Node 3 is Type II with $\oint \frac{\partial u}{\partial n} \phi_i \, ds = 0.5 \pm 0.3$ and
Nodes 4, 5, 6 are uniformly spaced Type I nodes
with 0.6 ± 0.2 and correlation length = 1 element
size

Then

$$\{\bar{b}\}_I = \begin{Bmatrix} 0 \\ 1.0 \\ 0 \\ 0 \\ \vdots \end{Bmatrix} ; \quad \{\bar{b}\}_N = \begin{Bmatrix} 0 \\ 0 \\ 0.5 \\ 0 \\ 0 \\ \vdots \end{Bmatrix} ; \quad \{\bar{b}\}_D = \begin{Bmatrix} 0 \\ 0 \\ 0 \\ 0.6 \\ 0.6 \\ 0.6 \\ 0 \\ \vdots \end{Bmatrix}$$

$$\left[Cov(\tilde{b})\right]_I = (0.5)^2 \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & \\ \vdots & \vdots & & & \ddots \end{bmatrix}$$

$$\left[Cov(\tilde{b})\right]_N = (0.3)^2 \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & \\ \vdots & \vdots & \vdots & & & \ddots \end{bmatrix}$$

$$\left[Cov(\tilde{b})\right]_D = (0.2)^2 \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & a & b & 0 & \cdots \\ 0 & 0 & 0 & a & 1 & a & 0 & \cdots \\ 0 & 0 & 0 & & & & & \\ 0 & 0 & 0 & b & a & 1 & 0 & \cdots \\ & & & & & & & \ddots \end{bmatrix}$$

$a \equiv 2e^{-1}$
$b \equiv 3e^{-2}$

- Noise Amplification: Is matrix inversion a noise amplifier?

<u>Special Case</u>: $[K]$ is square, real, symmetric

then $[K]$ has N independent eigenvalues/eigenvectors

i.e. $[K]\{V\}_i = \lambda_i \{V\}_i$

λ_i's are <u>real</u>

V_i's are orthogonal

$$\{V\}_i^T \{V\}_j = 0 \quad i \neq j$$
$$= 1 \quad i = j$$

Write $[V]$ as having columns $\{V\}_j$

then $[K][V] = [V][\mathrm{diag}(\lambda)]$

$$[K]\underbrace{[V][V]^T}_{I} = [V][\mathrm{diag}(\lambda)][V]^T$$

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

so $[K]^{-1} = [V][\mathrm{diag}(1/\lambda)][V]^T$ $\left\{ \begin{array}{l} \text{Since} \\ [V]^T[V] = I \\ [V][V]^T = I \end{array} \right.$

provided $\lambda_i \neq 0$ $i = 1, N$ (if any $\lambda_i = 0$, $[K]$ is singular, i.e. $[K]^{-1}$ does not exist) and for $[K]\{u\}=\{b\}$

$$\{u\}=[V]\left[\operatorname{diag}(1/\lambda)\right][V]^T\{b\}=\sum_{i=1}^{N}\{V_i\}\frac{\{V_i\}\cdot\{b\}}{\lambda_i}$$

But since $\{V\}_i$ are a complete basis for any N-dimensional vector can write

$$\{u\}=[V]\{c\} \iff \{c\}=[V]^T\{u\}$$

$$\{b\}=[V]\{d\} \iff \{d\}=[V]^T\{b\}$$

$\{c\}$ = projection of $\{u\}$ onto V-space

$\{d\}$ = projection of $\{b\}$ onto V-space

i.e. $\{u\}=\sum_i c_i\{V\}_i \iff c_i=\{V\}_i^T\{u\}$

$\{b\}=\sum_{i=1} d_i\{V\}_i \iff d_i=\{V\}_i^T\{b\}$

so $$\underbrace{[V]\{c\}}_{\{u\}}=[V]\left[\operatorname{diag}(1/\lambda)\right]\underbrace{[V]^T[V]}_{I}\underbrace{\{d\}}_{}$$

$$\{c\}=\left[\operatorname{diag}(1/\lambda)\right]\{d\}$$

or $c_i=\frac{1}{\lambda_i}d_i$

- If any $\lambda_i = 0$, then inversion is undefined
  <u>unless</u> $\{b\}^T\{V\}_i = 0$ (i.e. b is orthogonal to $\{V\}_i$)

- If λ_i small, noise projected onto $\{V\}_i$ through $\{d\}$ is amplified by $1/\lambda_i$ in $\{u\}$

- Inversion: $\{b\}$ is projected onto V-space, multiplied by $1/\lambda$ and the results projected back into $\{u\}$

<u>Criterion for noise amplification = small λ</u>

If only some λ_i small, inversion will amplify those components of $\{b\}$ selectively, i.e. part of $\{b\}$ which projects onto the associated $\{V\}_i$

- <u>Condition Number, K</u>
  - Equals ratio of largest to smallest λ in absolute value
  - $K = O(1)$, noise in $\{b\}$ passes to $\{u\}$ unchanged
  - $K >> 1$, noise is filtered and selectively amplified

$$[Cov(u)] = [V][diag(1/\lambda)][V]^T Cov(b) [V][diag(1/\lambda)][V]^T$$

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

If $[Cov(b)] = \sigma^2 I$, then $[Cov(d)] = \sigma^2 I$ and

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

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