Schur Transformation, General Case

Schur basis is NOT UNIQUE. But, sum of squares of FF weights = $\text{Tr}\,\mathbf{W}\mathbf{W}^\dagger - \sum_i |\lambda_i|^2$ is a unitary invariant. One natural scalar measure is this, or this over $\text{Tr}\,\mathbf{W}\mathbf{W}^\dagger$, or squareroot of either.

More generally: in Schur basis $\mathbf{W}-\Lambda = \mathbf{N}$ is upper triangular element; $\min_{\text{Schur decompositions}} |\mathbf{N}| =$ Henrici's "departure from normality" dep($\mathbf{W}$); $|\mathbf{N}|$ same for all decompositions for Frobenious norm. Another measure: $|\mathbf{W}\mathbf{W}^\dagger - \mathbf{W}^\dagger\mathbf{W}|$.

Trefethan & Embree:

> In the end, all scalar measures of nonnormality suffer from a basic limitation: Nonnormality is too complex to be summarized in a single number. Even when a matrix is 'highly nonnormal', one must know the gemoetry of that nonnormality in order to judge whether it will have important consequences for a given application."

For $\mathbf{J}$ random IID matrix, elements variance $1/N$: $\text{Tr}\,\mathbf{J}\mathbf{J}^\dagger = N^2\frac{1}{N} = N$. Sum over eigenvalues

$$= N \times \langle|\lambda^2|\rangle \tag{1}$$

$$= N\frac{\int_0^1 drr^3 \int d\theta}{\int_0^1 drr \int d\theta} \tag{2}$$

$$= N/2 \tag{3}$$

Example: submatrices simultaneously diagonizable by unitary transform

e.g. if translation invariant: $\rightarrow 2\times 2$ submatrices for each freq.

Example: Rajan and Abbott

Example: $\begin{pmatrix} \mathbf{W}_E & -\mathbf{W}_I \\ \mathbf{W}_E & -\mathbf{W}_I \end{pmatrix}$ go to sum and diff basis of those eigs; sum modes have eigs $\lambda_i^E - \lambda_i^I$; also N zero eigs; diff modes give $\lambda_i^E + \lambda_i^I$ times sum modes.

SLIDE

Schur Basis, 2×2 case

Assume two distinct eigenvectors, $\mathbf{e}_1$, e_2; complete e_1 with Schur basis vector $\mathbf{q}$:

$$\mathbf{q} = \frac{\mathbf{e}_2 - (\mathbf{e}_1\cdot\mathbf{e}_2)\mathbf{e}_1}{\sqrt{1-|\mathbf{e}_1\cdot\mathbf{e}_2|^2}} \tag{4}$$

Then compute

$$\mathbf{W}\mathbf{q} = \frac{\lambda_2 \mathbf{e}_2 - \lambda_1(\mathbf{e}_1 \cdot \mathbf{e}_2)\mathbf{e}_1}{\sqrt{1 - |\mathbf{e}_1 \cdot \mathbf{e}_2|^2}} = \lambda_2 \mathbf{q} + \frac{(\lambda_2 - \lambda_1)(\mathbf{e}_1 \cdot \mathbf{e}_2)}{\sqrt{1 - |\mathbf{e}_1 \cdot \mathbf{e}_2|^2}} \mathbf{e}_1 \tag{5}$$

Thus, in the Schur basis $(\mathbf{e}_1, \mathbf{q})$, $\mathbf{W}$ takes the upper triangular form

$$\begin{aligned} \mathbf{W} &= \begin{pmatrix} \lambda_1 & \beta \\ 0 & \lambda^+ \end{pmatrix} \text{ with} \\ \beta &= \frac{(\lambda_2 - \lambda_1)(\mathbf{e}_1 \cdot \mathbf{e}_2)}{\sqrt{1 - |\mathbf{e}_1 \cdot \mathbf{e}_2|^2}} \end{aligned} \tag{6}$$

β is

- Large when $|\mathbf{e}_1 \cdot \mathbf{e}_2|$ is close to one, i.e. when the angle between the eigenvectors is small;
- Zero when the matrix is normal, so that $|\mathbf{e}_1 \cdot \mathbf{e}_2| = 0$.

(It also becomes zero if $\lambda_1 = \lambda_2$, but this means that the matrix is normal, assuming there are two distinct eigenvectors.)

For biological weight matrix: for both eigenvectors, real parts of both elements are same sign: sum mode; $\mathbf{q}$ is difference mode[1].

Compute β for weight matrix:

- Eigs real: $x_+^2 - w_{EI} w_{IE} > 0$

$$\begin{aligned} |\beta|^2 &= \mathrm{Tr}\, \mathbf{W}\mathbf{W}^\dagger - |\lambda_+|^2 - |\lambda_-|^2 \qquad (10) \\ &= w_{EE}^2 + w_{EI}^2 + w_{IE}^2 + w_{II}^2 - 2(x_-^2 + x_+^2 - w_{EI} w_{IE}) \qquad (11) \\ &= (w_{EI} + w_{IE})^2 \qquad (12) \end{aligned}$$

  Note $\beta > 0$, therefore $\beta = w_{EI} + w_{IE}$. With real eigs, β is large when $\mathbf{W}$ deviates strongly from Hermitian.

[1] Eigenvalues:

$$\begin{aligned} \lambda_\pm &= \frac{w_{EE} - w_{II}}{2} \pm \sqrt{\left(\frac{w_{EE} + w_{II}}{2}\right)^2 - w_{EI} w_{IE}} \qquad (7) \\ &\equiv x_- \pm \sqrt{x_+^2 - w_{EI} w_{IE}} \qquad (8) \end{aligned}$$

Eigenvectors:

$$\mathbf{e}_\pm \propto \begin{pmatrix} 1 \\ (\lambda_\mp + w_{II})/w_{EI} \end{pmatrix} = \begin{pmatrix} 1 \\ \left(x_+ \mp \sqrt{x_+^2 - w_{EI} w_{IE}}\right)/w_{EI} \end{pmatrix} \tag{9}$$

Take $|\mathbf{e}_\pm| = 1$. Note: both elements have real parts > 0; sum mode. Assume eigenvectors are distinct.

Let Schur basis be $\{\mathbf{e}_-, \mathbf{q}\}$. $\mathbf{q} \propto \begin{pmatrix} (\lambda_+^* + w_{II})/w_{EI} \\ -1 \end{pmatrix}$: difference mode.

- Eigs complex: $w_{EI}w_{IE} - x_+^2 > 0$

$$\begin{aligned} |\beta|^2 &= \mathrm{Tr}\,\mathbf{W}\mathbf{W}^\dagger - |\lambda_+|^2 - |\lambda_-|^2 && (13)\\ &= w_{EE}^2 + w_{EI}^2 + w_{IE}^2 + w_{II}^2 - 2(x_-^2 + w_{EI}w_{IE} - x_+^2) && (14)\\ &= (w_{EI} - w_{IE})^2 + (w_{EE} + w_{II})^2 && (15) \end{aligned}$$

With complex eigs, β is large when $\mathbf{W}$ deviates strongly from anti-Hermitian.

Pseudospectra

Spectra: $\exists \mathbf{v}$ s.t. $\mathbf{W}\mathbf{v} = z\mathbf{v}$ for scalar z. In terms of resolvent of $\mathbf{W}$, $R_{\mathbf{W}}(z) = (W - z\mathbf{1})^{-1}$, spectrum is $\sigma(\mathbf{W}) = \{z : ||R_{\mathbf{W}}(z)|| = \infty\}$

ϵ-Pseudospectra: $\sigma_\epsilon(\mathbf{W}) = \{z : ||R_{\mathbf{W}}(z)|| > \frac{1}{\epsilon}\}$, *i.e.* for vector-induced norm, $\exists \mathbf{v}$ s.t. $|\mathbf{W}\mathbf{v} - z\mathbf{v}| < \epsilon$.

For vector-induced norm: equivalent definition is $\sigma_\epsilon(\mathbf{W}) = \{z : \exists \mathbf{d}W, ||\mathbf{d}W|| < \epsilon$ s.t. $z \in \sigma(\mathbf{W} + \mathbf{d}W)\}$. For normal matrix, pseudospectrum is just the set of ϵ-balls around the spectrum. For non-normal, it always includes that but is generally larger, can be much larger – spectrum very non-robust to perturbation.

Example: $\mathbf{M} = \begin{pmatrix} \lambda_1 & \beta \\ 0 & \lambda_2 \end{pmatrix}$. Perturb with $\begin{pmatrix} 0 & 0 \\ \epsilon & 0 \end{pmatrix}$, size ϵ. Spectrum is perturbed from $\{\lambda_1, \lambda_2\}$ to

$$\begin{aligned} & \frac{\lambda_1 + \lambda_2}{2} \pm \sqrt{\left(\frac{\lambda_1 - \lambda_2}{2}\right)^2 + \beta\epsilon} && (16)\\ = \;& \frac{\lambda_1 + \lambda_2}{2} \pm \frac{\lambda_1 - \lambda_2}{2}\sqrt{1 + \frac{4\beta\epsilon}{(\lambda_1 - \lambda_2)^2}} && (17)\\ = \;& \left\{\lambda_1 + \frac{\lambda_1 - \lambda_2}{2}\left(\sqrt{1 + \frac{4\beta\epsilon}{(\lambda_1 - \lambda_2)^2}} - 1\right),\right. && (18)\\ & \left.\lambda_2 - \frac{\lambda_1 - \lambda_2}{2}\left(\sqrt{1 + \frac{4\beta\epsilon}{(\lambda_1 - \lambda_2)^2}} - 1\right)\right\} && (19)\\ \to_{\beta\to\infty} \;& \frac{\lambda_1 + \lambda_2}{2} \pm \sqrt{\frac{\beta}{\epsilon}\epsilon} && (20) \end{aligned}$$

SLIDE

Kreiss Bound, bounds on transients

Pseudospectra give a lower bound on the size of transients of a dynamics $\frac{d\mathbf{x}}{dt} = \mathbf{W}\mathbf{x}$, where all the eigenvalues of $\mathbf{W}$ have negative real part:

$$\sup_{t\geq 0} ||e^{t\mathbf{W}}|| \geq \alpha_\epsilon(\mathbf{W})/\epsilon \quad \text{for all } \epsilon > 0 \tag{21}$$

where $\alpha_\epsilon(\mathbf{W})$ is the largest real part of an element of the ϵ-pseudospectrum of $\mathbf{W}$. This leads directly to the Kreiss bound: defining

$$\mathcal{K}(\mathbf{W}) = \sup_{\epsilon>0} \alpha_\epsilon(\mathbf{W})/\epsilon = \sup_{\Re z>0} (\Re(z))||(z-\mathbf{W})^{-1}|| \tag{22}$$

then

$$\sup_{t\geq 0} ||e^{t\mathbf{W}}|| \geq \mathcal{K}(\mathbf{W}) \tag{23}$$

also

$||e^{t\mathbf{W}} \leq eN\mathcal{K}(\mathbf{W})$

Other bounds:

Numerical range (or field of values) $W(\mathbf{W})$: the set $\{\mathbf{x}^\dagger\mathbf{W}\mathbf{x}$ for $|\mathbf{x}|=1$; equals eigenvalues of symmetric part of $\mathbf{W}$, $(\mathbf{W}+\mathbf{W}^\dagger)/2$. Contains convex hull of spectrum; = convex hull for normal matrix. Numerical abscissa $\omega(\mathbf{W})$ = maximum real part of numerical range = max eig of $(\mathbf{W}+\mathbf{W}^\dagger)/2$.

$\alpha(\mathbf{W})$: spectral abscissa: max real part of spectrum

For all t, $e^{t\alpha(\mathbf{W})} \leq ||e^{t\mathbf{W}}|| \leq e^{t\omega(\mathbf{W})}$

$||e^{t\mathbf{W}}|| \leq e^{t\omega(\mathbf{W})} \qquad \forall t$

$||e^{t\mathbf{W}}|| \leq e^{t\omega(\mathbf{W})} + o(t)$ as $t \to 0$

Proof of Kreiss Bound:

The bound in Eq. 21 is established as follows. There is always some real ω and $M \geq 1$ such that $||e^{t\mathbf{W}}|| \leq Me^{\omega t}$ for all $t \geq 0$. ω can be any number larger than the largest real part of an eigenvalue of $\mathbf{W}$ (you may have to choose M large to compensate for transients at finite times), so when all eigenvalues have negative real part we can always choose $\omega < 0$.

Then, for any z with $\Re z > w$, and in particular for any z with $\Re z > 0$,

$$(z-\mathbf{W})^{-1} = \int_0^\infty dt\, e^{-t(z-\mathbf{W})} = \int_0^\infty dt\, e^{-tz}e^{-t\mathbf{W}} \tag{24}$$

This implies

$$||(z-\mathbf{W})^{-1}|| \leq \int_0^\infty dt\, |e^{-tz}|\, ||e^{-t\mathbf{W}}|| \tag{25}$$

Let $S = \sup_{t\geq 0} ||e^{t\mathbf{W}}||$. Then

$$||(z-\mathbf{W})^{-1}|| \leq S\int_0^\infty dt\, |e^{-tz}| = S/(\Re(z)) \tag{26}$$

or

$$\sup_{t\geq 0} ||e^{t\mathbf{W}}|| \geq (\Re(z))||(z-\mathbf{W})^{-1}|| \tag{27}$$

How long do pseudospectra act like spectra?

Pseudospectrum can "act like eigenvalues" during transient period, if move to positive real part by more than ϵ. If z is in $\sigma_\epsilon(\mathbf{W})$, $\Re(z) > 0$, $\Re(z)/\epsilon > 1$, then for any $\tau > 0$,

$$\sup_{0<t\leq\tau} ||e^{tW}|| \geq e^{\Re(z)\tau} / \left(1 + \epsilon \frac{e^{\Re(z)\tau} - 1}{\Re(z)}\right) \tag{28}$$

Expression in paren is close to 1 for $e^{\Re(z)\tau} \ll 1 + \Re(z)/\epsilon$, i.e. $\tau < \frac{\log(1+\Re(z)/\epsilon)}{\Re(z)}$. On those time scales, z acts like an eigenvalue.

Bounds on pseudospectra

Eigenvalue condition number: $\kappa(\lambda) = \frac{|\mathbf{l}||\mathbf{r}|}{\mathbf{l}^\dagger \mathbf{r}} = \frac{1}{\cos(\theta_{\mathbf{lr}})}$ where $\mathbf{l}$ and $\mathbf{r}$ are left and right eigenvectors associated with eigenvalues. When all eigenvalues distinct: then as $\epsilon \to 0$[2]

$$\sigma_\epsilon(\mathbf{W}) \subseteq \cup_{j=1}^{N} \left(\lambda_j + \Delta_{\epsilon\kappa(\lambda_j)+O(\epsilon^2)}\right) \tag{37}$$

and for all $\epsilon > 0$, for vector-induced norms,

$$\sigma_\epsilon(\mathbf{W}) \subseteq \cup_{j=1}^{N} \left(\lambda_j + \Delta_{\epsilon N\kappa(\lambda_j)}\right) \tag{38}$$

Pseudospectra and numerical range:
$\sigma_\epsilon(\mathbf{W}) \subseteq W(\mathbf{W}) + \Delta_\epsilon$
and
$W(\mathbf{W})$ = intersection of all closed half-planes $H \in$ complex plane satisfying
$\sigma_\epsilon(\mathbf{W}) \subseteq H + \Delta_\epsilon\ \forall\epsilon > 0$.

---

[2]**Proof for $\epsilon \to 0$:** Consider $\mathbf{W}(t) = \mathbf{W} + t\mathbf{E}$ where $||\mathbf{E}|| = 1$. Perturbs to $\lambda(t)$, $\mathbf{r}(t)$. Because λ is simple root of characteristic equation, can expand as Taylor series in t. Note $\mathbf{l}$ is orthogonal to all other right eigenvectors, so writing $\mathbf{r}(t) = \mathbf{r} + \mathbf{d}r(t)$ with $\mathbf{d}r$ equal to t times a linear combination of the other right eigenvectors, then $\mathbf{l}\cdot\mathbf{r}(t) = \mathbf{l}\cdot\mathbf{r}$. So then:

$$\begin{aligned} \mathbf{l}^\dagger\mathbf{W}(t)\mathbf{r}(t) = \lambda(t)\mathbf{l}^\dagger\mathbf{r} &= \mathbf{l}^\dagger\mathbf{W}\mathbf{r}(t) + t\mathbf{l}^\dagger\mathbf{E}\mathbf{r}(t) &(29)\\ &= \lambda\mathbf{l}^\dagger\mathbf{r} + t\mathbf{l}^\dagger\mathbf{E}\mathbf{r}(t) &(30)\end{aligned}$$

So

$$\begin{aligned} |\lambda - \lambda(t)| &= |t|\frac{|\mathbf{l}^\dagger\mathbf{E}\mathbf{r}(t)|}{|\mathbf{l}^\dagger\mathbf{r}|} &(31)\\ &\le |t|\frac{|\mathbf{l}^\dagger||\mathbf{r}(t)|}{|\mathbf{l}^\dagger\mathbf{r}|} &(32)\\ &= |t|\frac{|\mathbf{l}^\dagger||\mathbf{r}|}{|\mathbf{l}^\dagger\mathbf{r}|} + O(|t|^2) &(33)\\ &= |t|\kappa(\lambda) + O(|t|^2) &(34)\end{aligned}$$

Proof for general ϵ: For each eig, normalize such that $\mathbf{l}^\dagger\mathbf{r} = 1$. Define $\mathbf{P}_j = \mathbf{r}_j\mathbf{l}_j^\dagger$; projection, $\mathbf{P}^2 = \mathbf{P}$. Note

$$||\mathbf{P}|| = \frac{||\mathbf{r}\mathbf{l}^\dagger||}{|\mathbf{l}^\dagger\mathbf{r}|} = \max_{|\mathbf{x}=1|}\frac{|\mathbf{r}\mathbf{l}^\dagger\mathbf{x}|}{|\mathbf{l}^\dagger\mathbf{r}|} = \frac{|\mathbf{r}||\mathbf{l}|}{|\mathbf{l}^\dagger\mathbf{r}|} = \kappa(\lambda) \tag{35}$$

where chose $\mathbf{x} = \mathbf{l}/|\mathbf{l}|$.

$z \in \sigma_\epsilon(\mathbf{W})$ implies $\exists\mathbf{x}$ with $|x| = 1$, $\mathbf{E}$ with $||\mathbf{E}|| < \epsilon$, s.t. $(\mathbf{W} + \mathbf{E})\mathbf{x} = z\mathbf{x}$. Then $\mathbf{P}\mathbf{E}\mathbf{x} = \mathbf{P}(z - \mathbf{W})\mathbf{x} = (z - \lambda)\mathbf{P}\mathbf{x}$. Provided $\mathbf{l}^\dagger\mathbf{x} \neq 0$, this implies for any eigenvalue, eigenvector that

$$|\lambda - z| = \frac{|\mathbf{P}\mathbf{E}\mathbf{x}|}{\mathbf{P}\mathbf{x}} < \epsilon\frac{||\mathbf{P}||}{|\mathbf{P}\mathbf{x}|} \tag{36}$$

But $1 = |\mathbf{x}| = |\sum_j \mathbf{P}_j\mathbf{x}| \le \sum_j |\mathbf{P}_j\mathbf{x}|$, so there must be at least one j with $|\mathbf{P}_j\mathbf{x}| > 1/N$. For this j, $|\lambda_j - z| < \epsilon N||\mathbf{P}_j|| = \epsilon N\kappa(\lambda_j)$ as required.