Appendix 1 – Stability

We calculate the Jacobian matrix $Df_{\mathbf{w}}$, for a fixed vector $\mathbf{w}$:

Lemma 0.1. $Df_{\mathbf{w}} = \mathbf{I} + \gamma\left[\mathbf{C} - 2\mathbf{w}(\mathbf{C}\mathbf{w})^T - (\mathbf{w}^T\mathbf{C}\mathbf{w})\mathbf{I}\right]$

Proof. Call $g(\mathbf{w}) = (\mathbf{w}^T\mathbf{C}\mathbf{w})\mathbf{w}$, so $f(\mathbf{w}) = \mathbf{w} + \gamma(\mathbf{C}\mathbf{w} - g(\mathbf{w}))$

$$g_i(\mathbf{w}) = (\mathbf{w}^T\mathbf{C}\mathbf{w})w_i$$

If $i \neq j$:

$$\frac{\partial g_i}{\partial w_j}(\mathbf{w}) = \frac{\partial}{\partial w_j}(\sum_{k,l} C_{kl}w_k w_l)w_i = 2(\sum_k C_{kj}w_k)w_i \;=\; 2[\mathbf{C}\mathbf{w}]_j w_i$$

If $i = j$:

$$\frac{\partial g_i}{\partial w_i}(\mathbf{w}) = \frac{\partial}{\partial w_i}(\sum_{k,l} C_{kl}w_k w_l)w_i + \sum_{k,l} C_{kl}w_k w_l = 2(\sum_k C_{ki}w_k)w_i +$$

$$+\mathbf{w}^T\mathbf{C}\mathbf{w} = 2[\mathbf{C}\mathbf{w}]_i w_i + \mathbf{w}^T\mathbf{C}\mathbf{w}$$

So:

$$Dg_{\mathbf{w}} = 2\mathbf{w}(\mathbf{C}\mathbf{w})^T + (\mathbf{w}^T\mathbf{C}\mathbf{w})\mathbf{I}$$

□

Take now an orthonormal basis $\mathcal{B}$ of eigenvectors of $\mathbf{C}$ (with respect to the Euclidean norm $\|\cdot\|$ on $\mathbb{R}^n$). Fix a vector $\mathbf{w} \in \mathcal{B}$. Pick any $\mathbf{v} \in \mathcal{B}, \mathbf{v} \neq \mathbf{w}$. Call $\lambda_{\mathbf{w}}$ and $\lambda_{\mathbf{v}}$ their corresponding eigenvalues.

$$\begin{aligned} Df_{\mathbf{w}}(\mathbf{v}) &= v + \gamma[\mathbf{C}\mathbf{v} - 2\mathbf{w}(\mathbf{C}\mathbf{w})^T\mathbf{v} - (\mathbf{w}^T\mathbf{C}\mathbf{w})\mathbf{v}] = \\ &= \mathbf{v} + \gamma[\mathbf{C}\mathbf{v} - 2\mathbf{w}\mathbf{w}^T\mathbf{C}\mathbf{v} - (\mathbf{w}^T\mathbf{C}\mathbf{w})\mathbf{w}] = \\ &= v + \gamma[\lambda_{\mathbf{v}}\mathbf{v} - 2\mathbf{w}\mathbf{w}^T\lambda_{\mathbf{v}}\mathbf{v} - \lambda_{\mathbf{w}}\mathbf{w}] = (1 - \gamma[\lambda_{\mathbf{w}} - \lambda_{\mathbf{v}}])\mathbf{v} \end{aligned}$$

$$\begin{aligned} Df_{\mathbf{w}}(\mathbf{w}) &= \mathbf{w} + \gamma[\mathbf{C}\mathbf{w} - 2\mathbf{w}(\mathbf{C}\mathbf{w})^T\mathbf{w} - (\mathbf{w}^T\mathbf{C}\mathbf{w})] = \\ &= \mathbf{w} + \gamma[\lambda_{\mathbf{w}}\mathbf{w} - 2\mathbf{w}\mathbf{w}^T\lambda_{\mathbf{w}}\mathbf{w} - \lambda_{\mathbf{w}}\mathbf{w}] = \\ &= \mathbf{w} + \gamma[-2\lambda_{\mathbf{w}}\|\mathbf{w}\|\mathbf{w}] = [1 - 2\gamma\lambda_{\mathbf{w}}]\mathbf{w} \end{aligned}$$

So $\mathcal{B}$ is also a basis of eigenvectors for $Df_{\mathbf{w}}$.

Our next goal is to generalize this argument for an iteration function that includes errors. The new model introduces an error matrix, $\mathbf{E} \in \mathcal{M}_n(\mathbb{R})$ that has positive entries, is symmetric and equal to the identity matrix $I \in \mathcal{M}_n(\mathbb{R})$ in case the error is zero. Moreover, we assume that $\mathbf{E}\mathbf{C}$ has strictly positive maximal eigenvalue of multiplicity one.

$$f^{\mathbf{E}}(\mathbf{w}) = \mathbf{w} + \gamma[\mathbf{E}\mathbf{C}\mathbf{w} - (\mathbf{w}^T\mathbf{C}\mathbf{w})\mathbf{w}]$$

Note that the symmetric, positive definite matrix $\mathbf{C} \in \mathcal{M}_n(\mathbb{R})$ defines a dot product in $\mathbb{R}^n$ as:

$$\langle \mathbf{v}, \mathbf{w} \rangle_{\mathbf{C}} = \mathbf{v}^T \mathbf{C} \mathbf{w}$$

If $\mathbf{v}$ and $\mathbf{w}$ are eigenvectors of $\mathbf{EC}$ corresponding to the eigenvalues $\lambda_{\mathbf{v}} \neq \lambda_{\mathbf{w}}$, then they are orthogonal with respect to the dot product $\langle , \rangle_{\mathbf{C}}$. Indeed:

$$\mathbf{ECv} = \lambda_{\mathbf{v}} \mathbf{v} \;\Rightarrow\; \langle \mathbf{w}, \mathbf{ECv} \rangle_{\mathbf{C}} = \lambda_{\mathbf{v}} \langle \mathbf{w}, \mathbf{v} \rangle_{\mathbf{C}}$$

$$\mathbf{ECw} = \lambda_{\mathbf{w}} \mathbf{w} \;\Rightarrow\; \langle \mathbf{v}, \mathbf{ECw} \rangle_{\mathbf{C}} = \lambda_{\mathbf{w}} \langle \mathbf{v}, \mathbf{w} \rangle_{\mathbf{C}}$$

Hence $\lambda_{\mathbf{v}} \langle \mathbf{v}, \mathbf{w} \rangle_{\mathbf{C}} = \lambda_{\mathbf{w}} \langle \mathbf{v}, \mathbf{w} \rangle_{\mathbf{C}}$. As $\lambda_{\mathbf{v}} \neq \lambda_{\mathbf{w}}$, it follows that $\langle \mathbf{v}, \mathbf{w} \rangle_{\mathbf{C}} = 0$, hence $\mathbf{v}$ and $\mathbf{w}$ are orthogonal with respect to the given dot product.

A fixed point for $f^{\mathbf{E}}$ is a vector $\mathbf{w} = (w_1 ... w_n)^T$ such that $\mathbf{ECw} = (\mathbf{w}^T \mathbf{Cw})\mathbf{w}$. In other words, $\mathbf{w}$ is fixed by $f^{\mathbf{E}}$ if and only if it is an eigenvector of $\mathbf{EC}$ (with corresponding eigenvalue $\lambda_{\mathbf{w}}$), normalized such that $\parallel \mathbf{w} \parallel_{\mathbf{C}} = \lambda_{\mathbf{w}}$. Clearly, this is possible if and only if $\lambda_{\mathbf{w}} > 0$.

$$\mathbf{ECw} = \lambda_{\mathbf{w}} \mathbf{w}, \quad \|\mathbf{w}\|_{\mathbf{C}} = \lambda_{\mathbf{w}}$$

If the multiplicity of $\lambda_{\mathbf{w}}$ is one, then $\mathbf{w}$ is orthogonal in $\langle , \rangle_{\mathbf{C}}$ to all other eigenvectors of $\mathbf{EC}$.

Recall that

$$Df^{\mathbf{E}}_{\mathbf{w}} = \mathbf{I} + \gamma[\mathbf{EC} - 2\mathbf{w}(\mathbf{Cw})^T - (\mathbf{w}^T \mathbf{Cw})\mathbf{I}]$$

Take $\mathbf{w}$ to be a fixed point of $f^{\mathbf{E}}$. $\mathbf{w}$ will hence be an eigenvector of $\mathbf{EC}$, with eigenvalue $\lambda_{\mathbf{w}} = (\mathbf{w}^T \mathbf{Cw})\mathbf{w} > 0$. Calculate:

$$\begin{aligned} Df^{\mathbf{E}}_{\mathbf{w}} \mathbf{w} &= \mathbf{w} + \gamma[\mathbf{ECw} - 2\mathbf{w}(\mathbf{Cw})^T \mathbf{w} - (\mathbf{w}^T \mathbf{Cw})\mathbf{w}] = \\ &= \mathbf{w} + \gamma[-2\mathbf{w}\mathbf{w}^T \mathbf{Cw}] = [1 - 2\gamma\lambda_{\mathbf{w}}]\mathbf{w} \end{aligned}$$

$$\begin{aligned} Df^{\mathbf{E}}_{\mathbf{w}} \mathbf{v} &= \mathbf{v} + \gamma[\mathbf{ECv} - 2\mathbf{w}\mathbf{w}^T \mathbf{Cv} - \lambda_{\mathbf{w}} \mathbf{v}] = \\ &= \mathbf{v} + \gamma[(\lambda_{\mathbf{v}} - \lambda_{\mathbf{w}})v - 2\langle \mathbf{w}, \mathbf{v} \rangle_{\mathbf{C}} \mathbf{w}] = (1 - \gamma[\lambda_{\mathbf{w}} - \lambda_{\mathbf{v}}])\mathbf{v} \end{aligned}$$

for any other eigenvector $\mathbf{v}$ of $\mathbf{EC}$ with eigenvalue $\lambda_{\mathbf{v}} \neq \lambda_{\mathbf{w}}$:

As in the error free case, $Df^{\mathbf{E}}_{\mathbf{w}}$ has all eigenvalues less than one in absolute value if and only if $\lambda_{\mathbf{w}}$ is the principal eigenvalue of $\mathbf{EC}$ and $\gamma < \frac{1}{\lambda_{\mathbf{w}}}$.

To complete our mathematical discussion, we argue that the condition on $\mathbf{EC}$ to have a unique maximal eigenvalue (i.e. a principal eigenspace of dimension one) is not unrealistically strict, and that $\mathbf{EC}$ has this property generically often. We will prove a somewhat stronger result. We start by assuming, without loss of genericity, that $\mathbf{C}$ has distinct eigenvalues $\lambda_1 > \lambda_2 > ... > \lambda_n > 0$. We will show that, via some assumptions which are also generically true, $\mathbf{EC}$ has also distinct eigenvalues.

$$\mathbf{E} = \begin{pmatrix} Q & \epsilon & \cdot & \epsilon \\ \epsilon & Q & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \epsilon & \cdot & \epsilon & Q \end{pmatrix} = (Q - \epsilon)\mathbf{I} + [\boldsymbol{\epsilon}] = (1 - n\epsilon)\mathbf{I} + [\boldsymbol{\epsilon}]$$

where $[\boldsymbol{\epsilon}]$ is the $n \times n$ matrix with all entries ϵ. We take $0 \leq \epsilon < 1/n$ (i.e., positive error ϵ smaller than the trivial value). We want to see what conditions need to be met in order to have a unique maximal eigenvalue for the modified covariance matrix $\mathbf{EC}$.

Since $\mathbf{C}$ is symmetric and positive semidefinite, there exists an orthogonal matrix $\mathbf{P}$ (i.e., $\mathbf{PP}^T = \mathbf{P}^T\mathbf{P} = \mathbf{I}$), so that $\mathbf{PCP}^T = \mathbf{D}$, where $\mathbf{D}$ is a diagonal matrix. Without loosing genericity, we can assume that $\mathbf{P}$ is such that the sum of the entries along each of its rows is nonzero.

Under there assumptions, we can proceed to calculate the eigenvalues of $\mathbf{PECP}^T$, since these will be the same as the eigenvalues of $\mathbf{EC}$. (Indeed, if $\mathbf{v}$ is an eigenvector of with eigenvalue μ, then $\mathbf{P}^T\mathbf{v}$ is an eigenvector of $\mathbf{EC}$, and conversely.) To do this, we first simplify the form of $\mathbf{PECP}^T$, then calculate its characteristic polynomial:

$$\mathbf{PECP}^T = \mathbf{PEP}^T\mathbf{PCP}^T = \Big((1-n\epsilon)\mathbf{I} + \mathbf{P}[\boldsymbol{\epsilon}]\mathbf{P}^T\Big)\,\mathbf{D}$$

For each $j = \overline{1,n}$, we call $r_j = \Sigma_{k=1}^n p_{kj}$ the sum of the entries on the j-th row of $\mathbf{P}$. With this notation:

$$\begin{aligned}
\mathbf{PECP}^T &= \left[(1-n\epsilon)I + \epsilon \begin{pmatrix} r_1^2 & r_1r_2 & \cdot & r_1r_n \\ r_2r_1 & r_2^2 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ r_nr_1 & \cdot & \cdot & r_n^2 \end{pmatrix}\right] \begin{pmatrix} \lambda_1 & 0 & \cdot & 0 \\ 0 & \lambda_2 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & \lambda_n \end{pmatrix} \\
&= (1-n\epsilon) \begin{pmatrix} \lambda_1(ar_1^2+1) & \lambda_2 ar_1r_2 & \cdot & \lambda_n ar_1r_n \\ \lambda_1 ar_2r_1 & \lambda_2(ar_2^2+1) & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \lambda_1 ar_nr_1 & \cdot & \cdot & \lambda_n(ar_n^2+1) \end{pmatrix}
\end{aligned}$$

where $a = \dfrac{\epsilon}{1-n\epsilon}$, for $0 < \epsilon < 1/n$.

The eigenvalues of $\mathbf{PECP}^T$ (which are also the eigenvalues of $\mathbf{EC}$ are proportional to the roots of the polynomial $P(x) = \det\left(\dfrac{1}{1-n\epsilon}\mathbf{PECP}^T - x\mathbf{I}\right)$:

$$\begin{aligned}
P(x) &= \begin{vmatrix} \lambda_1(ar_1^2+1)-x & \lambda_2 ar_1r_2 & \cdot & \lambda_n ar_1r_n \\ \lambda_1 ar_2r_1 & \lambda_2(ar_2^2+1)-x & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \lambda_1 ar_nr_1 & \cdot & \cdot & \lambda_n(ar_n^2+1)-x \end{vmatrix} \\
&= a^n\lambda_1\lambda_2...\lambda_n r_1^2r_2^2...r_n^2 \begin{vmatrix} 1+\frac{\lambda_1-x}{a\lambda_1r_1^2} & 1 & \cdot & 1 \\ 1 & 1+\frac{\lambda_2-x}{a\lambda_2r_2^2} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & 1+\frac{\lambda_n-x}{a\lambda_nr_n^2} \end{vmatrix}
\end{aligned}$$

We want to calculate $P(\lambda_j)$, for all the distinct eigenvalues λ_j of $\mathbf{C}$. We describe the computation for λ_n, since the others are very similar.

$$P(\lambda_n) = a^n\lambda_1\lambda_2...\lambda_n r_1^2r_2^2...r_n^2 \begin{vmatrix} 1+\frac{\lambda_1-\lambda_n}{a\lambda_1r_1^2} & 1 & \cdot & 1 \\ 1 & 1+\frac{\lambda_2-\lambda_n}{a\lambda_2r_2^2} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & 1 \end{vmatrix}$$

We subtract the last row from all others, and expand along the n-th column:

$$\begin{aligned} P(\lambda_n) &= a^n \lambda_1 \lambda_2 ... \lambda_n r_1^2 r_2^2 ... r_n^2 \begin{vmatrix} \frac{\lambda_1-\lambda_n}{a\lambda_1 r_1^2} & 0 & \cdot & 0 \\ 0 & \frac{\lambda_2-\lambda_n}{a\lambda_2 r_2^2} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & 1 \end{vmatrix} \\ &= a^n \lambda_1 \lambda_2 ... \lambda_n r_1^2 r_2^2 ... r_n^2 (-1)^{2n} \cdot \frac{\lambda_1 - \lambda_n}{a\lambda_1 r_1^2} \cdot \frac{\lambda_2 - \lambda_n}{a\lambda_2 r_2^2} ... \frac{\lambda_{n-1} - \lambda_n}{a\lambda_{n-1} r_{n-1}^2} \\ &= a\lambda_n r_n^2 (\lambda_1 - \lambda_n)(\lambda_2 - \lambda_n)...(\lambda_{n-1} - \lambda_n) > 0 \end{aligned}$$

We obtain similar expressions for the other $P(\lambda_j)$, for $\lambda_1 > \lambda_2 > ... > \lambda_n$ and conclude that $\text{sign}(P(\lambda_j)) = (-1)^{n-j}$, hence $P(x)$ has a root between each two consecutive eigenvalues of $\mathbf{C}$ (these are $n-1$ of the n-th roots of P). Moreover, since $P(\lambda_1) = (-1)^n$ and $\lim_{x\to\infty} P(x) = (-1)^n \infty$, then the n-th and largest root of P is between λ_1 and ∞.
We have therefore obtained more than our desired conclusion: $P(x)$ has n positive, distinct roots, hence the matrix $\mathbf{EC}$ has n distinct positive real eigenvalues. In particular, it has a unique maximal eigenvalue. This concludes the proof.