We remember the transfer operator T,

$$[Tf](y) = \frac{1}{\mu(y)} \int\limits_{\Omega} \mathrm{d}x\, p(x,y)\mu(x)f(x),$$

and its spectral decomposition,

$$Tf = \sum_{i=1}^{\infty} \lambda_m \langle f, \psi_m \rangle_\mu \psi_m,$$

with

$$T\psi_m = \lambda_m \psi_m \quad \text{and} \quad \langle \psi_m, \psi_{m\prime} \rangle_\mu = \delta_{m,m\prime}.$$

The variational principle

Theorem: The scalar product of Tf and f has an upper bound given by

$$\langle Tf, f \rangle_\mu \leq \begin{cases} 1, & \langle f, f \rangle_\mu = 1 \\ \lambda_k, & \langle f, f \rangle_\mu = 1,\, \langle f, \psi_i \rangle_\mu = 0\, \forall i < k. \end{cases} \tag{1}$$

Proof. We coose a $k > 1$ and expand f in the basis of the transfer operator's eigenvectors,

$$f = \sum_{i=1}^{\infty} c_i \psi_i, \quad \text{with} \quad c_i = 0\, \forall i < k.$$

Thus, the scalar product reads

$$\begin{aligned} \langle Tf, f \rangle_\mu &= \sum_{i,j=k}^{\infty} c_i c_j \langle T\psi_i, \psi_j \rangle_\mu \\ &= \sum_{i,j=k}^{\infty} c_i c_j \lambda_i \langle \psi_i, \psi_j \rangle_\mu \\ &= \sum_{i=k}^{\infty} c_i^2 \lambda_i \\ &\leq \lambda_k \sum_{i=k}^{\infty} c_i^2 = \lambda_k. \end{aligned}$$

□

The idea of the variational principle is to choose a basis $\chi_1, \dots, \chi_N$ and find a linear combination

$$f(\mathbf{b}) = \sum_{i=1}^{N} b_i \chi_i \tag{2}$$

which satisfies $\|f(\mathbf{b})\|_2 = 1$ and maximises the scalar product $\langle Tf(\mathbf{b}), f(\mathbf{b})\rangle_\mu$ over the set of all possible linear combinations.

Theorem: The vector $\mathbf{b}$, with $\langle Tf(\mathbf{b}), f(\mathbf{b})\rangle_\mu = \max$, is determined by the eigenvector to the largest eigenvalue of the generalised eigenvalue problem

$$C^\tau \mathbf{b} = \hat{\lambda} C^0 \mathbf{b}, \tag{3}$$

with

$$C^\tau_{i,j} = \langle T\chi_i, \chi_j\rangle_\mu \quad \text{and} \quad C^0_{i,j} = \langle \chi_i, \chi_j\rangle_\mu . \tag{4}$$

Proof. We choose a suitable cost function,

$$\begin{aligned} L((\mathbf{b})) &= \langle Tf(\mathbf{b}), f(\mathbf{b})\rangle_\mu - \xi \left(\langle f(\mathbf{b}), f(\mathbf{b})\rangle_\mu - 1 \right) \\ &= \sum_{i,j=1}^{N} b_i b_j \langle T\chi_i, \chi_j\rangle_\mu - \xi \left(\sum_{i,j=1}^{N} b_i b_j \langle \chi_i, \chi_j\rangle_\mu - 1 \right), \end{aligned}$$

and set its derivative w.r.t. the components of $(\mathbf{b})$ to zero:

$$0 \stackrel{!}{=} \frac{\partial L}{\partial b_{i'}} = 2\sum_{j=1}^{N} b_j \langle T\chi_{i'}, \chi_j\rangle_\mu - 2\xi \sum_{j=1}^{N} b_j \langle \chi_{i'}, \chi_j\rangle_\mu$$

$$\Leftrightarrow \quad \sum_{j=1}^{N} b_j \langle T\chi_{i'}, \chi_j\rangle_\mu = \xi \sum_{j=1}^{N} b_j \langle \chi_{i'}, \chi_j\rangle_\mu$$

$$\Leftrightarrow \quad \sum_{j=1}^{N} C^\tau_{i',j} b_j = \xi \sum_{j=1}^{N} C^0_{i',j} b_j$$

□

The relation of C^τ and C^0 to MSMs and TICA

Remembering the integral expressions from last week's lecture, we can rewrite

$$C^0_{i,j} = \langle \chi_i, \chi_j\rangle_\mu = \int_\Omega \mathrm{d}x\, \mu(x)\chi_i(x)\chi_j(x) \tag{5}$$

and

$$
\begin{aligned}
C^{\tau}_{i,j} &= \langle T\chi_i, \chi_j \rangle_\mu \\
&= \int\limits_\Omega \mathrm{d}y\, \mu(y) \left(T\chi_i\right)(y)\chi_j(y) \\
&= \int\limits_\Omega \mathrm{d}y\, \mu(y) \frac{1}{\mu(y)} \int\limits_\Omega \mathrm{d}x\, p(x,y)\mu(x)\chi_i(x)\chi_j(y) \\
&= \int\limits_\Omega \mathrm{d}y \int\limits_\Omega \mathrm{d}x\, \mathbb{P}\left(X_0 = x, X_1 = y\right) \chi_i(x)\chi_j(y).
\end{aligned}
$$

To understand the relation between C^τ, C^0 and Markov state models, we define a decomposition of the configuration space,

$$
\Omega = \bigcup_{i=1}^{N} \mathcal{S}_i,
$$

with

$$
\mathcal{S}_i \cap \mathcal{S}_j = \emptyset \quad \text{if } i \neq j,
$$

and choose indicator functions as basis:

$$
\chi_i(x) = \begin{cases} 1, & x \in \mathcal{S}_i \\ 0, & x \notin \mathcal{S}_i. \end{cases} \tag{6}
$$

By substituting this choice into the integral expressions, we find

$$
\begin{aligned}
C^{0}_{i,j} &= \int\limits_\Omega \mathrm{d}x\, \mu(x)\chi_i(x)\chi_j(x) \\
&= \delta_{i,j} \int\limits_\Omega \mathrm{d}x\, \mu(x)\chi_i(x) \\
&= \delta_{i,j} \int\limits_{\mathcal{S}_i} \mathrm{d}x\, \mu(x) \\
&= \delta_{i,j}\pi_i
\end{aligned}
$$

and

$$
\begin{aligned}
C^{0}_{i,j} &= \int\limits_\Omega \mathrm{d}y \int\limits_\Omega \mathrm{d}x\, \mathbb{P}\left(X_0 = x, X_1 = y\right) \chi_i(x)\chi_j(y) \\
&= \int\limits_{\mathcal{S}_j} \mathrm{d}y \int\limits_{\mathcal{S}_i} \mathrm{d}x\, \mathbb{P}\left(X_0 = x, X_1 = y\right) \\
&= \mathbb{P}\left(X_0 \in \mathcal{S}_i, X_1 \in \mathcal{S}_j\right).
\end{aligned}
$$

Therefore, $\left(C^0\right)^{-1}$ exists and we can write

$$\left(\left(C^0\right)^{-1} C^\tau\right)_{i,j} = \frac{C^\tau_{i,j}}{C^0_{i,j}} = \frac{\mathbb{P}\left(X_0 \in \mathcal{S}_i, X_1 \in \mathcal{S}_j\right)}{\mathbb{P}\left(X_0 \in \mathcal{S}_i\right)} = \mathbb{P}\left(X_1 \in \mathcal{S}_j | X_0 \in \mathcal{S}_i\right), \quad (7)$$

and we have recovered a transition matrix from C^τ and C^0 via the variational principle.

The time-lagged independent component analysis (TICA) is a method for dimensionality reduction. In case of TICA, we choose the basis of the configuration space Ω as $\chi_1, \ldots, \chi_N$ and sample the matrices C^τ and C^0 from trajectory data:

$$\begin{aligned} C^0_{i,j} &= \int_\Omega \mathrm{d}x\, \mu(x)\chi_i(x)\chi_j(x) \\ &\approx \frac{1}{M}\sum_{t=1}^{M} \chi_i(x_t)\chi_j(x_t) \end{aligned}$$

$$\begin{aligned} C^\tau_{i,j} &= \int_\Omega \mathrm{d}y \int_\Omega \mathrm{d}x\, \mathbb{P}\left(X_0 = x, X_1 = y\right)\chi_i(x)\chi_j(y) \\ &\approx \frac{1}{M-1}\sum_{t=1}^{M-1} \chi_i(x_t)\chi_j(x_{t+1}) \end{aligned}$$