Calculus for some operators on multivariate function space

Bo Markussen (bomar@math.ku.dk)
Department of Mathematical Sciences
University of Copenhagen

September 11, 2014
September 10, 2015: Small corrections made

1 Introduction

This manuscript describes an algebra of linear operators defined on the space $\mathcal{C}_{\mathrm{PW}}([a, b]; \mathbb{R}^{N \times M})$ of piecewise continuous functions defined on a compact interval $[a, b]$ with values in the space of $N \times M$-matrices. For brevity this space will be denoted by $\mathcal{H}$ in the following. The interval $[a, b]$ and an associated sampling mesh $a = t_0 = t_1 < \ldots < t_p = t_{p+1} = b$ with $p \geq 2$ will be fixed, whereas the dimensions $N, M \in \mathbb{N}$ may change. When necessary we write $\mathcal{H}(N, M)$ to specify the matrix dimensions. Furthermore, we will assume that functions $f \in \mathcal{H}$ are continuous from the right with limits from the left, and that possible points of discontinuity only may occur at t_j for $j = 1, \ldots, p - 1$.

2 Operator families

This section introduces different mathematical classes of linear operators on the function space $\mathcal{H}$. These classes have been called *families* in the section title to avoid confusion with the `R` classes introduced later.

2.1 Pointwise operators

The transpose $f^\top$ and the sum $f + g$ of $f, g \in \mathcal{H}(N, M)$, and the matrix product $f * g$ of $f \in \mathcal{H}(N, K)$ and $g \in \mathcal{H}(K, M)$ are defined pointwise. Furthermore, for $f \in \mathcal{H}(N, M_1)$ and $g \in \mathcal{H}(N, M_2)$ the concatenation $f|g \in$

$\mathcal{H}(N, M_1 + M_2)$ is defined by $(f|g)(t) = \big(f(t), g(t)\big)$. And as is common we sometimes will omit the asterisk in the notation for the matrix product.

2.2 Elementary operators

The *forward integral* F, the *backward integral* B, the *integral* I, and the *triangular integrals* T are defined on

$$\mathrm{F}\colon \mathcal{H}(N,M) \to \mathcal{H}(N,M), \qquad \mathrm{B}\colon \mathcal{H}(N,M) \to \mathcal{H}(N,M),$$
$$\mathrm{I}\colon \mathcal{H}(N,M) \to \mathbb{R}^{N\times M}, \qquad \mathrm{T}\colon \mathcal{H}(N,M) \to \mathbb{R}^{N\times M\times p}$$

and for $f \in \mathcal{H}(N,M)$ and $t \in [a,b]$ given by

$$\mathrm{F}f(t) = \int_a^t f(s)\,\mathrm{d}s, \qquad \mathrm{B}f(t) = \int_t^b f(s)\,\mathrm{d}s, \qquad \mathrm{I}f = \int_a^b f(s)\,\mathrm{d}s,$$

and

$$\mathrm{T}f = \left\{ \int_{t_{j-1}}^{t_j} \frac{t - t_{j-1}}{t_j - t_{j-1}} f(t)\,\mathrm{d}t + \int_{t_j}^{t_{j+1}} \frac{t_{j+1} - t}{t_{j+1} - t_j} f(t)\,\mathrm{d}t \right\}_{j=1,\dots,p}.$$

Note that interchanging the order of integration gives identities like

$$\mathrm{I}\big(f\mathrm{F}(g)\big) = \mathrm{I}\big(\mathrm{B}(f)g\big), \qquad \mathrm{I}\big(f\mathrm{B}(g)\big) = \mathrm{I}\big(\mathrm{F}(f)g\big).$$

2.3 Multiplication operators

Multiplication operators will in general be denoted by the letter $\mathcal{M}$. A multiplication operator $\mathcal{M}$ is described by a function $\alpha \in \mathcal{H}(N,N)$, and its action is given by pointwise multiplication

$$\mathcal{M}f = \alpha * f, \quad f \in \mathcal{H}. \tag{1}$$

The identity operator on $\mathcal{H}$ is denoted by $\mathcal{I}$, and $\mathcal{I}$ is a multiplication operator.

2.4 Triangle operators

Triangle operators will in general be denoted by the letter $\mathcal{T}$. A triangle operator $\mathcal{T}$ is described by two functions $\beta, \gamma \in \mathcal{H}(N,K)$ for some $K \in \mathbb{N}$, and its action is given by

$$\mathcal{T}f = \beta * \mathrm{F}(\gamma^\top * f), \quad f \in \mathcal{H}, \tag{2}$$

i.e. $\mathcal{T}f(t) = \beta(t)\int_a^t \gamma(s)^\top f(s)\,\mathrm{d}s$.

2.5 Projection operators

Projection operators will in general be denoted by the letter $\mathcal{P}$. A projection operator $\mathcal{P}$ is described by two functions $\delta, \epsilon \in \mathcal{H}(N, L)$ for some $L \in \mathbb{N}$, and its action is given by

$$\mathcal{P}f = \delta * \mathrm{I}(\epsilon^\top * f), \quad f \in \mathcal{H}, \tag{3}$$

i.e. $\mathcal{P}f(t) = \delta(t) \int_a^b \epsilon(s)^\top f(s)\,\mathrm{d}s$.

2.6 Lattice operators

Lattice operators will in general be denoted by the letter $\mathcal{L}$. A lattice operator $\mathcal{L} = \mathcal{M} + \mathcal{T} + \mathcal{P}$ is given as a sum of a multiplication, a triangle, and a projection operator. We will also call such a sum for a lattice operator when the multiplication part vanishes. The 5-function parametrization of a lattice operator

$$\mathcal{L}f = \alpha * f + \beta * \mathrm{F}(\gamma^\top * f) + \delta * \mathrm{I}(\epsilon^\top * f), \quad f \in \mathcal{H}, \tag{4}$$

is given by the quintet $(\alpha, \beta, \gamma, \delta, \epsilon)$.

2.7 Symmetric lattice operators

Symmetric lattice operators will in general be denote by the letter $\mathcal{S}$. The 5-function parametrization of a symmetric lattice operator $\mathcal{S} \sim (\alpha, \beta, \gamma, \delta, \epsilon)$ has $\alpha = \alpha^\top$, and we may assume the form

$$\beta = \epsilon| - \delta, \qquad\qquad \gamma = \delta|\epsilon$$

with $\delta, \epsilon \in \mathcal{H}(N, L)$ and $K = 2L$. In particular, we have a 3-functions parametrization of a symmetric lattice operator given by the triplet $(\alpha, \delta, \epsilon)$.

3 Algebraic properties

The following table displays the structural result of adding two operators from the families introduced in sections 2.3 to 2.6:

Sum $A_1 + A_2$	Term A_2:			
Term A_1:	$\mathcal{M}$	$\mathcal{T}$	$\mathcal{P}$	$\mathcal{L}$
$\mathcal{M}$	$\mathcal{M}$	$\mathcal{M}+\mathcal{T}$	$\mathcal{M}+\mathcal{P}$	$\mathcal{L}$
$\mathcal{T}$	$\mathcal{M}+\mathcal{T}$	$\mathcal{T}$	$\mathcal{L}$	$\mathcal{L}$
$\mathcal{P}$	$\mathcal{M}+\mathcal{P}$	$\mathcal{L}$	$\mathcal{P}$	$\mathcal{L}$
$\mathcal{L}$	$\mathcal{L}$	$\mathcal{L}$	$\mathcal{L}$	$\mathcal{L}$

The following table displays the structural result of multiplying two operators from the classes introduced above:

Product $A_1 * A_2$	Factor A_2:			
Factor A_1:	$\mathcal{M}$	$\mathcal{T}$	$\mathcal{P}$	$\mathcal{L}$
$\mathcal{M}$	$\mathcal{M}$	$\mathcal{T}$	$\mathcal{P}$	$\mathcal{L}$
$\mathcal{T}$	$\mathcal{T}$	$\mathcal{T}$	$\mathcal{P}$	$\mathcal{L}$
$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{L}$
$\mathcal{L}$	$\mathcal{L}$	$\mathcal{L}$	$\mathcal{L}$	$\mathcal{L}$

The following table displays the structural result of inverting an invertible operator from the classes introduced above:

Operator A	$\mathcal{M}$	$\mathcal{M}+\mathcal{T}$	$\mathcal{M}+\mathcal{P}$	$\mathcal{L}$
Inverse A^{-1}	$\mathcal{M}$	$\mathcal{M}+\mathcal{T}$	$\mathcal{M}+\mathcal{P}$	$\mathcal{L}$

Moreover, the sum of two symmetric lattice operators is a symmetric lattice operator, and the inverse of a symmetric lattice operator is a symmetric lattice operator.

3.1 Factorization

The $\mathcal{TP}$-factorization of a lattice operator $\mathcal{L} = \mathcal{M} + \mathcal{T} + \mathcal{P}$ is given by

$$\mathcal{M} + \mathcal{T} + \mathcal{P} = (\mathcal{M} + \mathcal{T})(\mathcal{I} + \mathcal{P}_*), \qquad \mathcal{P}_* = (\mathcal{M} + \mathcal{T})^{-1}\mathcal{P},$$

and the $\mathcal{PT}$-factorization is given by

$$\mathcal{M} + \mathcal{T} + \mathcal{P} = (\mathcal{I} + \mathcal{P}_\star)(\mathcal{M} + \mathcal{T}), \qquad \mathcal{P}_\star = \mathcal{P}(\mathcal{M} + \mathcal{T})^{-1}.$$

We note that $\mathcal{P}_*$ and $\mathcal{P}_\star$ are projection operators. Both factorizations may be used for the inversion of a lattice operator via inversion of the factors.

3.2 Addition of lattice operators

Using the 5-function parametrization the sum of two lattice operators is given by

$$(\alpha_1, \beta_1, \gamma_1, \delta_1, \epsilon_1) + (\alpha_2, \beta_2, \gamma_2, \delta_2, \epsilon_2) = (\alpha_1 + \alpha_2, \beta_1|\beta_2, \gamma_1|\gamma_2, \delta_1|\delta_2, \epsilon_1|\epsilon_2).$$

The (K, L)-dimension in the 5-function parametrization for the sum of two lattice operators is given by

$$(K_1, L_1) + (K_2, L_2) = (K_1 + K_2, L_1 + L_2).$$

3.3 Multiplication of lattice operators

Using the 5-function parametrization the product of two lattice operators is given

$$
\begin{aligned}
&(\alpha_1, \beta_1, \gamma_1, \delta_1, \epsilon_1) * (\alpha_2, \beta_2, \gamma_2, \delta_2, \epsilon_2) \\
&\qquad\qquad = \Big(\alpha_1\alpha_2, \\
&\qquad\qquad\quad \beta_1\big|\alpha_1\beta_2 - \beta_1\mathrm{B}(\gamma_1^\top\beta_2), \quad \alpha_2^\top\gamma_1 + \gamma_2\mathrm{B}(\beta_2^\top\gamma_1)\big|\gamma_2, \\
&\qquad\qquad\quad \delta_1\big|\alpha_1\delta_2 + \beta_1\mathrm{F}(\gamma_1^\top\delta_2), \quad \alpha_2^\top\epsilon_1 + \gamma_2\mathrm{B}(\beta_2^\top\epsilon_1) + \epsilon_2\mathrm{I}(\delta_2^\top\epsilon_1)\big|\epsilon_2\Big).
\end{aligned}
$$

The (K, L)-dimension in the 5-function parametrization for the product of two lattice operators is given by

$$(K_1, L_1) * (K_2, L_2) = (K_1 + K_2, L_1 + L_2).$$

The above results follow invoking linearity on the following table:

Product $A_1 * A_2$	Factor A_2:		
Factor A_1:	$(\alpha_2, 0, 0, 0, 0)$	$(0, \beta_2, \gamma_2, 0, 0)$	$(0, 0, 0, \delta_2, \epsilon_2)$
$(\alpha_1, 0, 0, 0, 0)$	$(\alpha_1\alpha_2, 0, 0, 0, 0)$	$(0, \alpha_1\beta_2, \gamma_2, 0, 0)$	$(0, 0, 0, \alpha_1\delta_2, \epsilon_2)$
$(0, \beta_1, \gamma_1, 0, 0)$	$(0, \beta_1, \alpha_2^\top\gamma_1, 0, 0)$	$\big(0, \beta_1\vert - \beta_1\mathrm{B}(\gamma_1^\top\beta_2), \gamma_2\mathrm{B}(\beta_2^\top\gamma_1)\vert\gamma_2, 0, 0\big)$	$\big(0, 0, 0, \beta_1\mathrm{F}(\gamma_1^\top\delta_2), \epsilon_2\big)$
$(0, 0, 0, \delta_1, \epsilon_1)$	$(0, 0, 0, \delta_1, \alpha_2^\top\epsilon_1)$	$\big(0, 0, 0, \delta_1, \gamma_2\mathrm{B}(\beta_2^\top\epsilon_1)\big)$	$\big(0, 0, 0, \delta_1, \epsilon_2\mathrm{I}(\delta_2^\top\epsilon_1)\big)$

Note also that the (K, L)-dimensions for multiplication are given by

Product $A_1 * A_2$	Factor A_2:	
Factor A_1:	$(K_2, 0)$	$(0, L_2)$
$(K_1, 0)$	$(K_1 + K_2, 0)$	$(0, L_2)$
$(0, L_1)$	$(0, L_1)$	$(0, L_1)$

3.4 Inversion

The solution to the matrix value Volterra integral equation gives the inversion

$$(\alpha, \beta, \gamma, 0, 0)^{-1} = (\alpha^{-1}, -\alpha^{-1}\beta\kappa^\top, \alpha^{-1,\top}\gamma\kappa^{-1}, 0, 0),$$

where $\kappa\colon [0, 1] \to \mathbb{R}^{K\times K}$ solves the linear differential equation

$$\kappa(0) = \mathbb{I}_K, \qquad \tfrac{\mathrm{d}}{\mathrm{d}t}\kappa(t) = -\kappa(t)\beta(t)^\top\alpha(t)^{-1,\top}\gamma(t).$$

And elementary matrix calculus gives the inversion

$$(\mathbb{I}_N, 0, 0, \delta, \epsilon)^{-1} = \Big(\mathbb{I}_N, 0, 0, -\delta\big(\mathbb{I}_L + \mathrm{I}(\epsilon^\top\delta)\big)^{-1}, \epsilon\Big).$$

Inversion of a general lattice operator is done applying the $\mathcal{PT}$-factorization

$$\big(\mathcal{M}+\mathcal{T}+\mathcal{P}\big)^{-1} = (\mathcal{M}+\mathcal{T})^{-1}(\mathcal{I}+\mathcal{P}_\star)^{-1}, \qquad \mathcal{P}_\star = \mathcal{P}(\mathcal{M}+\mathcal{T})^{-1}.$$

This gives

$$\begin{aligned}(\alpha,\beta,\gamma,\delta,\epsilon)^{-1} &= (\alpha^{-1},\tilde\beta,\tilde\gamma,0,0) * \Big(\mathbb{I}_N,0,0,-\delta\big(\mathbb{I}_L+\mathrm{I}(\epsilon_\star^\top\delta)\big)^{-1},\epsilon_\star\Big)\\ &= \Big(\alpha^{-1},\tilde\beta,\tilde\gamma,-\delta_\star\big(\mathbb{I}_L+\mathrm{I}(\epsilon_\star^\top\delta)\big)^{-1},\epsilon_\star\Big)\end{aligned}$$

with

$$\begin{aligned}\tilde\beta &= -\alpha^{-1}\beta\kappa^\top,\\ \tilde\gamma &= \alpha^{-1,\top}\gamma\kappa^{-1},\\ \delta_\star &= (\mathcal{M}+\mathcal{T})^{-1}\delta = \alpha^{-1}\delta+\tilde\beta * \mathrm{F}(\tilde\gamma^\top\delta),\\ \epsilon_\star &= \alpha^{-1,\top}\epsilon+\tilde\gamma * \mathrm{B}(\tilde\beta^\top\epsilon).\end{aligned} \tag{5}$$

Here the formula for $\epsilon_\star$ follows from

$$\begin{aligned}(0,0,0,\delta,\epsilon_\star) &= \mathcal{P}_\star = \mathcal{P}(\mathcal{M}+\mathcal{T})^{-1}\\ &= (0,0,0,\delta,\epsilon) * (\alpha^{-1},\tilde\beta,\tilde\gamma,0,0)\\ &= \big(0,0,0,\delta,\alpha^{-1,\top}\epsilon+\tilde\gamma * \mathrm{B}(\tilde\beta^\top\epsilon)\big).\end{aligned}$$

3.5 Inversion of symmetric operators

Let $\mathcal{S}=(\alpha,\delta,\epsilon)$ be a symmetric lattice operator, i.e. the associated 5-function parametrization satisfies

$$\alpha=\alpha^\top, \qquad \beta=\epsilon|-\delta, \qquad \gamma=\delta|\epsilon.$$

Then γ, δ, and ϵ may be inferred from β via

$$\gamma=\beta\begin{pmatrix}0 & \mathbb{I}_L\\ -\mathbb{I}_L & 0\end{pmatrix}, \qquad \delta=\beta\begin{pmatrix}0\\ -\mathbb{I}_L\end{pmatrix}, \qquad \epsilon=\beta\begin{pmatrix}\mathbb{I}_L\\ 0\end{pmatrix}.$$

Suppose $\mathcal{S}^{-1}$ is derived using Eq. (5). Then we have

$$\begin{aligned}\left\{\dot\kappa^\top\begin{pmatrix}0 & \mathbb{I}_L\\ -\mathbb{I}_L & 0\end{pmatrix}\right\} &= -\gamma^\top\alpha^{-1}\beta\kappa^\top\begin{pmatrix}0 & \mathbb{I}_L\\ -\mathbb{I}_L & 0\end{pmatrix}\\ &= \begin{pmatrix}0 & \mathbb{I}_L\\ -\mathbb{I}_L & 0\end{pmatrix}\beta^\top\alpha^{-1}\beta\left\{\kappa^\top\begin{pmatrix}0 & \mathbb{I}_L\\ -\mathbb{I}_L & 0\end{pmatrix}\right\}\end{aligned}$$

and

$$\left\{ \begin{pmatrix} 0 & \mathbb{I}_L \\ -\mathbb{I}_L & 0 \end{pmatrix} (\dot{\kappa^{-1}}) \right\} = \begin{pmatrix} 0 & \mathbb{I}_L \\ -\mathbb{I}_L & 0 \end{pmatrix} \beta^\top \alpha^{-1,\top} \gamma \kappa^{-1}$$
$$= \begin{pmatrix} 0 & \mathbb{I}_L \\ -\mathbb{I}_L & 0 \end{pmatrix} \beta^\top \alpha^{-1} \beta \left\{ \begin{pmatrix} 0 & \mathbb{I}_L \\ -\mathbb{I}_L & 0 \end{pmatrix} \kappa^{-1} \right\},$$

which together with $\kappa(0)^\top = \kappa(0)^{-1} = \mathbb{I}_K$ imply

$$\kappa^\top \begin{pmatrix} 0 & \mathbb{I}_L \\ -\mathbb{I}_L & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mathbb{I}_L \\ -\mathbb{I}_L & 0 \end{pmatrix} \kappa^{-1},$$

and hence

$$\tilde{\gamma} = \alpha^{-1,\top} \gamma \kappa^{-1} = \alpha^{-1} \beta \begin{pmatrix} 0 & \mathbb{I}_L \\ -\mathbb{I}_L & 0 \end{pmatrix} \kappa^{-1} = \alpha^{-1} \beta \kappa^\top \begin{pmatrix} 0 & \mathbb{I}_L \\ -\mathbb{I}_L & 0 \end{pmatrix}$$
$$= \tilde{\beta} \begin{pmatrix} 0 & \mathbb{I}_L \\ -\mathbb{I}_L & 0 \end{pmatrix}.$$

3.6 Numerical inversion of symmetric lattice operators

Let $\mathcal{S}_1 = (\alpha_1, \delta_1, \epsilon_1)$ and $\mathcal{S}_2 = (\alpha_2, \delta_2, \epsilon_2)$ be two symmetric lattice operators. Then we have

$$\mathcal{S}_1\mathcal{S}_2 = (\alpha_1\alpha_2, \beta_{12}, \gamma_{12}, \delta_1|\delta_{12}, \epsilon_{12}|\epsilon_2),$$
$$\mathcal{S}_2\mathcal{S}_1 = (\alpha_2\alpha_1, \beta_{21}, \gamma_{21}, \delta_2|\delta_{21}, \epsilon_{21}|\epsilon_1),$$

where $\beta_{12}, \beta_{21}, \gamma_{12}, \gamma_{21}$ are some functions, and where

$$\delta_{12} = \alpha_1\delta_2 + \epsilon_1 \mathrm{F}(\delta_1^\top \delta_2) - \delta_1 \mathrm{F}(\epsilon_1^\top \delta_2),$$
$$\delta_{21} = \alpha_2\delta_1 + \epsilon_2 \mathrm{F}(\delta_2^\top \delta_1) - \delta_2 \mathrm{F}(\epsilon_2^\top \delta_1),$$
$$\epsilon_{12} = \alpha_2^\top \epsilon_1 + \delta_2 \mathrm{B}(\epsilon_2^\top \epsilon_1) + \epsilon_2 \mathrm{F}(\delta_2^\top \epsilon_1),$$
$$\epsilon_{21} = \alpha_1^\top \epsilon_2 + \delta_1 \mathrm{B}(\epsilon_1^\top \epsilon_2) + \epsilon_1 \mathrm{F}(\delta_1^\top \epsilon_2).$$

This implies that a sufficient condition for $\mathcal{S}_1 = \mathcal{S}_2^{-1}$ is

$$\alpha_1\alpha_2 - \mathbb{I}_N = 0, \qquad \delta_{12} = 0, \qquad \delta_{21} = 0, \qquad \epsilon_{12} = 0, \qquad \epsilon_{21} = 0,$$

and this is also necessary if δ_1 and ϵ_2 are not collinear and if δ_2 and ϵ_1 are not collinear. Thus, the inverse operator may be found minimizing the functional

$$E = \tfrac{1}{2}\mathrm{I}\big(\mathrm{tr}\big((\alpha_1\alpha_2 - \mathbb{I}_N)^\top(\alpha_1\alpha_2 - \mathbb{I}_N)\big)\big)$$
$$+ \tfrac{1}{2}\mathrm{I}\big(\mathrm{tr}(\delta_{12}^\top\delta_{12})\big) + \tfrac{1}{2}\mathrm{I}\big(\mathrm{tr}(\delta_{21}^\top\delta_{21})\big) + \tfrac{1}{2}\mathrm{I}\big(\mathrm{tr}(\epsilon_{12}^\top\epsilon_{12})\big) + \tfrac{1}{2}\mathrm{I}\big(\mathrm{tr}(\epsilon_{21}^\top\epsilon_{21})\big).$$

The minimum equals 0, the functions α_1 and $\delta_1\epsilon_1^\top$ are unique at the minimizer, and we may compute the gradient. Using identities like $\mathrm{I}(f\mathrm{B}(g)) = \mathrm{I}(\mathrm{F}(f)g)$ we find

$$\frac{\partial E}{\partial \alpha_1} = \alpha_1\alpha_2\alpha_2^\top - \alpha_2^\top + \delta_{12}\delta_2^\top + \epsilon_2\epsilon_{21}^\top,$$
$$\frac{\partial E}{\partial \delta_1} = \delta_2\mathrm{B}(\delta_{12}^\top\epsilon_1) - \delta_{12}\mathrm{F}(\delta_2^\top\epsilon_1) + \alpha_2^\top\delta_{21} + \delta_2\mathrm{B}(\epsilon_2^\top\delta_{12}) - \epsilon_2\mathrm{B}(\delta_2^\top\delta_{12}) + \epsilon_{21}\mathrm{B}(\epsilon_2^\top\epsilon_1) + \epsilon_2\mathrm{B}(\epsilon_{21}^\top\epsilon_1)$$
$$\frac{\partial E}{\partial \epsilon_1} = \delta_{12}\mathrm{F}(\delta_2^\top\delta_1) - \delta_2\mathrm{B}(\delta_{12}^\top\delta_1) + \alpha_2\epsilon_{12} + \epsilon_2\mathrm{F}(\delta_2^\top\epsilon_{12}) + \delta_2\mathrm{B}(\epsilon_2^\top\epsilon_{12}) + \epsilon_2\mathrm{F}(\epsilon_{21}^\top\delta_1) + \epsilon_{12}\mathrm{F}(\epsilon_2^\top\delta_1)$$

and for later usage we also compute the derivative w.r.t. $\mathcal{S}_2$, i.e.

$$\frac{\partial E}{\partial \alpha_2} = \alpha_1^\top\alpha_1\alpha_2 - \alpha_1^\top + \delta_{21}\delta_1^\top + \epsilon_1\epsilon_{12}^\top,$$
$$\frac{\partial E}{\partial \delta_2} = \alpha_1^\top\delta_{12} + \delta_1\mathrm{B}(\epsilon_1^\top\delta_{12}) - \epsilon_1\mathrm{B}(\delta_1^\top\delta_{12}) + \delta_1\mathrm{B}(\delta_{21}^\top\epsilon_2) - \delta_{21}\mathrm{F}(\delta_1^\top\epsilon_2) + \epsilon_{12}\mathrm{B}(\epsilon_1^\top\epsilon_2) + \epsilon_1\mathrm{B}(\epsilon_{12}^\top\epsilon_2),$$
$$\frac{\partial E}{\partial \epsilon_2} = \delta_{12}\mathrm{F}(\delta_1^\top\delta_2) - \delta_1\mathrm{B}(\delta_{21}^\top\delta_2) + \epsilon_1\mathrm{F}(\epsilon_{12}^\top\delta_2) + \epsilon_{12}\mathrm{F}(\epsilon_1^\top\delta_2) + \alpha_1\epsilon_{21} + \epsilon_1\mathrm{F}(\delta_1^\top\epsilon_{21}) + \delta_1\mathrm{B}(\epsilon_1^\top\epsilon_{21}).$$

3.7 Dimension reduction (to be updated)

The formulae stated above may give 5-function parametrizations with unnecessarily large dimensions K or L. To investigate how these dimensions possibly may be reduced let us consider the triangular operator $\beta\langle\gamma 1_{(0,\cdot]}, \cdot\rangle = (0, \beta, \gamma, 0, 0)$. Let $UDU^\top \in \mathbb{R}^{K\times K}$ be the eigen decomposition of the positive semi-definite matrix $\langle\beta, \beta\rangle \in \mathbb{R}^{K\times K}$, and let $U_0 \in \mathbb{R}^{K\times K_0}$ be the columns of U corresponding to the non-zero eigenvalues. Then we have $\beta(t) = \beta(t)UU^\top = \beta(t)U_0U_0^\top$, and hence

$$(0, \beta, \gamma, 0, 0) = (0, \beta U_0, \gamma U_0, 0, 0)$$

with $\beta U_0, \gamma U_0 \colon [0, 1] \to \mathbb{R}^{q\times K_0}$ providing a dimension reduction if $K_0 < K$. The same eigenvalue analysis may be applied on the parameters γ, δ and ϵ.

4 Implementation in R

The mathematical objects introduced above have been implemented in an `R` package named `operatorCalc` using S4 classes.

4.1 Piecewise polynomial matrix functions

A function $f \in \mathcal{H}(N, M)$ is encoded in an S4 class named `matFct` with the following slots:

- **@`mesh`**: An object of class `mesh`, which is a strictly increasing numeric sequence of length at least two encoding the sampling mesh $a = t_1 < \ldots < t_p = b$.
- **@`f`**: (N, M, p)-array with the sampling of f at the knot points t_j.
- **@`g`**: $(N, M, p-1, r)$-array representing the polynomial interpolation of f in the intervals $[t_j, t_{j+1})$. Here $r = 0$ is also legal, and implies that the sum over k vanishes in Eq. (6)
- **@`continuous`** $(p-1)$-logical vector. The j'th entry specifies whether the function is assumed to be continuous on the interval $[t_j, t_{j+1}]$ in the approximation done when taking pointwise matrix inverses.

The evaluation of a `matFct` object x at $t \in [0, 1]$ is defined by

$$\texttt{eval}(t, x) = (1 - s) \cdot x@\texttt{f}[, , j] + s \cdot x@\texttt{f}[, , j+1] + \sum_{k=1}^{r} s^k \cdot x@\texttt{g}[, , j, k] \quad (6)$$

when $t \in [t_j, t_{j+1})$ and with $s = \frac{t - t_j}{t_{j+1} - t_j}$, and where $t = b$ is included in the closed interval $[t_{p-1}, t_p]$. We see that $\texttt{eval}(t_j, x) = x@\texttt{f}[, , j]$ for $j = 1, \ldots, p-1$. Futhermore, $\texttt{eval}(\cdot, x)$ is continuous if and only if

$$\sum_{k=1}^{r} x@\texttt{g}[, , , k] = 0.$$

In the following subsections we describe the arithmetic methods defined on the `matFct` class.

4.1.1 Sum of matrix functions

The sum of two `matFct` objects x and y with the same sampling mesh and congruent matrix dimensions is given by

$$(x + y)@\texttt{f} = x@\texttt{f} + y@\texttt{f},$$
$$(x + y)@\texttt{g} = x@\texttt{g} + y@\texttt{g},$$

where the terms in the g-slot are padded with zeros if the polynomial orders are different. A possible way to do this is to define $q = \min\{x@\texttt{r}, y@\texttt{r}\}$, $r = \max\{x@\texttt{r}, y@\texttt{r}\}$, and if $q > 1$ then

$$(x + y)@\texttt{g}[, , , 1 : (q-1)] = x@\texttt{g}[, , , 1 : (q-1)] + y@\texttt{g}[, , , 1 : (q-1)],$$

and if $q < r$ then

$$(x+y)@\mathtt{g}[,,q:(r-1)] = \begin{cases} x@\mathtt{g}[,,q:(r-1)] & \text{if } r = x@\mathtt{r}, \\ y@\mathtt{g}[,,q:(r-1)] & \text{if } r = y@\mathtt{r}. \end{cases}$$

4.1.2 Product of matrix functions

The sum of two `matFct` objects x and y with the same sampling mesh and congruent matrix dimensions is given by $(x*y)@\mathtt{r} = x@\mathtt{r} + y@\mathtt{r}$ and

$$(x*y)@\mathtt{f}[,,j] = x@\mathtt{f}[,,j] * y@\mathtt{f}[,,j],$$

and $(x*y)@\mathtt{g}[,,j,h]$ is given as the coefficient of s^h on the right hand side of

$$\begin{aligned} x(t)*y(t) &- (1-s)\cdot x@\mathtt{f}[,,j]*y@\mathtt{f}[,,j] - s\cdot x@\mathtt{f}[,,j+1]*y@\mathtt{f}[,,j+1] \\ &= (-s+s^2)\cdot \big(x@\mathtt{f}[,,j+1] - x@\mathtt{f}[,,j]\big) * \big(y@\mathtt{f}[,,j+1] - y@\mathtt{f}[,,j]\big) \\ &\quad + \sum_{k=1}^{y@\mathtt{r}} s^k \cdot x@\mathtt{f}[,,j]*y@\mathtt{g}[,,j,k] + \sum_{k=1}^{y@\mathtt{r}} s^{k+1}\cdot \big(x@\mathtt{f}[,,j+1]-x@\mathtt{f}[,,j]\big)*y@\mathtt{g}[,,j,k] \\ &\quad + \sum_{k=1}^{x@\mathtt{r}} s^k \cdot x@\mathtt{g}[,,j,k]*y@\mathtt{f}[,,j] + \sum_{k=1}^{x@\mathtt{r}} s^{k+1}\cdot x@\mathtt{g}[,,j,k]*\big(y@\mathtt{f}[,,j+1]-y@\mathtt{f}[,,j]\big) \\ &\quad + \sum_{k=1}^{x@\mathtt{r}}\sum_{l=1}^{y@\mathtt{r}} s^{k+l}\cdot x@\mathtt{g}[,,j,k]*y@\mathtt{g}[,,j,l]. \end{aligned}$$

4.1.3 Forward integral of a matrix function

The forward integral $\int_0^t x(u)\,\mathrm{d}u$ of a `matFct` object x is given by

$$\begin{aligned} \mathrm{F}x@\mathtt{f}[,,1] &= 0, \\ \mathrm{F}x@\mathtt{f}[,,j+1] &\overset{j\geq 1}{=} \mathrm{F}x@\mathtt{f}[,,j] + (t_{j+1}-t_j)\cdot\left(\frac{x@\mathtt{f}[,,j]+x@\mathtt{f}[,,j+1]}{2} + \sum_{k=1}^{x@\mathtt{r}} \frac{x@\mathtt{g}[,,j,k]}{k+1}\right), \\ \mathrm{F}x@\mathtt{g}[,,j,2] &= (t_{j+1}-t_j)\cdot\frac{x@\mathtt{f}[,,j+1]-x@\mathtt{f}[,,j]+x@\mathtt{g}[,,j,1]}{2}, \\ \mathrm{F}x@\mathtt{g}[,,j,k] &\overset{3\leq k\leq r+1}{=} (t_{j+1}-t_j)\cdot\frac{x@\mathtt{g}[,,j,k-1]}{k}, \\ \mathrm{F}x@\mathtt{g}[,,j,1] &= -\sum_{k=2}^{r+1}\mathrm{F}x@\mathtt{g}[,,j,k]. \end{aligned}$$

4.1.4 Backward integral of a matrix function

The backward integral $\int_t^1 x(u)\,\mathrm{d}u = \int_{t_j}^1 x(u)\,\mathrm{d}u - \int_{t_j}^t x(u)\,\mathrm{d}u$ of a `matFct` object x is given by

$$\begin{aligned}
\mathrm{B}x@\mathtt{f}[,,p] &= 0, \\
\mathrm{B}x@\mathtt{f}[,,j-1] &\overset{j\le p}{=} \mathrm{B}x@\mathtt{f}[,,j] + (t_j - t_{j-1}) \cdot \left(\frac{x@\mathtt{f}[,,j-1] + x@\mathtt{f}[,,j]}{2} + \sum_{k=1}^{x@\mathtt{r}} \frac{x@\mathtt{g}[,,j-1,k]}{k+1} \right), \\
\mathrm{B}x@\mathtt{g}[,,j,2] &= -(t_{j+1} - t_j) \cdot \frac{x@\mathtt{f}[,,j+1] - x@\mathtt{f}[,,j] + x@\mathtt{g}[,,j,1]}{2}, \\
\mathrm{B}x@\mathtt{g}[,,j,k] &\overset{3\le k\le r+1}{=} -(t_{j+1} - t_j) \cdot \frac{x@\mathtt{g}[,,j,k-1]}{k}, \\
\mathrm{B}x@\mathtt{g}[,,j,1] &= -\sum_{k=2}^{r+1} \mathrm{B}x@\mathtt{g}[,,j,k].
\end{aligned}$$

4.1.5 Triangular evaluation of a matrix function

The triangular evaluation $\left\{ \int_{t_{j-1}}^{t_j} \frac{t-t_{j-1}}{t_j-t_{j-1}} x(t)\,\mathrm{d}t + \int_{t_j}^{t_{j+1}} \frac{t_{j+1}-t}{t_{j+1}-t_j} x(t)\,\mathrm{d}t \right\}_{j=1,\dots,p}$ of a `matFct` object x is an (N,M,p)-array Tx with

$$\begin{aligned}
\mathrm{T}x[,,1] &= (t_2 - t_1) \cdot \left(\frac{x@\mathtt{f}[,,1]}{3} + \frac{x@\mathtt{f}[,,2]}{6} + \sum_{k=1}^{x@\mathtt{r}} \frac{x@\mathtt{g}[,,1,k]}{(k+1)(k+2)} \right) \\
\mathrm{T}x[,,j] &\overset{1<j<p}{=} (t_j - t_{j-1}) \cdot \left(\frac{x@\mathtt{f}[,,j-1]}{6} + \frac{x@\mathtt{f}[,,j]}{3} + \sum_{k=1}^{x@\mathtt{r}} \frac{x@\mathtt{g}[,,j-1,k]}{k+2} \right) \\
&\qquad + (t_{j+1} - t_j) \cdot \left(\frac{x@\mathtt{f}[,,j]}{3} + \frac{x@\mathtt{f}[,,j+1]}{6} + \sum_{k=1}^{x@\mathtt{r}} \frac{x@\mathtt{g}[,,j,k]}{(k+1)(k+2)} \right) \\
\mathrm{T}x[,,p] &= (t_p - t_{p-1}) \cdot \left(\frac{x@\mathtt{f}[,,p-1]}{6} + \frac{x@\mathtt{f}[,,p]}{3} + \sum_{k=1}^{x@\mathtt{r}} \frac{x@\mathtt{g}[,,p-1,k]}{k+2} \right),
\end{aligned}$$

where $t_0 = 0$ and $t_{p+1} = 1$.

4.1.6 Super sampling

Below we describe the recoding of a `matFct` object x sampled at a finer mesh. Suppose that $t_j \le t_n^{\mathrm{new}} < t_{n+1}^{\mathrm{new}} \le t_{j+1}$, and let $t \in [t_n^{\mathrm{new}}, t_{n+1}^{\mathrm{new}})$ be given. Then

$$s = \frac{t - t_j}{t_{j+1} - t_j} = \frac{t_{n+1}^{\mathrm{new}} - t_n^{\mathrm{new}}}{t_{j+1} - t_j} s_{\mathrm{new}} + \frac{t_n^{\mathrm{new}} - t_j}{t_{j+1} - t_j}, \qquad s_{\mathrm{new}} = \frac{t - t_n^{\mathrm{new}}}{t_{n+1}^{\mathrm{new}} - t_n^{\mathrm{new}}},$$

and hence

$$\begin{aligned}\sum_{l=1}^{r} g_l \cdot s^l &= \sum_{l=1}^{r}\sum_{k=0}^{l} \binom{l}{k} \left(\frac{t_n^{\text{new}} - t_j}{t_{j+1} - t_j}\right)^{l-k} \left(\frac{t_{n+1}^{\text{new}} - t_n^{\text{new}}}{t_{j+1} - t_j}\right)^{k} g_l \cdot s_{\text{new}}^k \\ &= \sum_{l=1}^{r} \left(\frac{t_n^{\text{new}} - t_j}{t_{j+1} - t_j}\right)^{l} g_l + \sum_{k=1}^{r}\sum_{l=k}^{r} \binom{l}{k} \left(\frac{t_n^{\text{new}} - t_j}{t_{j+1} - t_j}\right)^{l-k} \left(\frac{t_{n+1}^{\text{new}} - t_n^{\text{new}}}{t_{j+1} - t_j}\right)^{k} g_l \cdot s_{\text{new}}^k.\end{aligned}$$

Here the first term should be added to the f^{new} of the super sampled function, the second term contributes to g_k^{new}, and the linear part added via $\text{diff}(f^{\text{new}})$ should be removed from the g_1^{new}.

4.2 Lattice operator

A lattice operator $\mathcal{L}$ is encoded in an S4 class named `operator` with the following slots:

@alpha: `matFct` object.

@beta: `matFct` object.

@gamma: `matFct` object.

@delta: `matFct` object.

@epsilon: `matFct` object.

4.3 Symmetric lattice operator

A symmetric lattice operator $\mathcal{S}$ is encoded in an S4 class named `symmOperator` with the following slots:

@alpha: `matFct` object.

@delta: `matFct` object.

@epsilon: `matFct` object.

In principle the alpha-sloth should contain a symmetric matrix valued function. This is not checked, but the alpha-sloth will be symmetrized in the operator inversion.

5 Likelihood estimation of $\mathcal{S}$ operators

Let $X_1, \dots, X_n$ be i.id. from $\mathcal{N}(0, \hat{\mathcal{S}}_0)$ equidistant sampled at $t_j = \frac{j-1}{p-1}$ for $j = 1, \dots, p$, where $\mathcal{S}_0$ is some symmetric lattice operator and the covariance matrix $\hat{\mathcal{S}}_0 \in \mathbb{R}^{p\times p}$ is defined below. Let $(\alpha, \delta, \epsilon)$ be the 3-function parametrization of symmetric lattice operators with some $N, L \in \mathbb{N}$, and let the upper triangular part of the symmetric matrix $\hat{\mathcal{S}}_{\alpha,\delta,\epsilon} \in \mathbb{R}^{(N\times N)\times(p\times p)}$ be defined by

$$\hat{\mathcal{S}}_{\alpha,\delta,\epsilon}(i,j) = \alpha(t_j) 1_{i=j} + p^{-1}\delta(t_i)\epsilon(t_j)^\top \in \mathbb{R}^{N\times N}, \quad \text{for } 1 \le i \le j \le p.$$

The inverse of a symmetric lattice operator is a symmetric lattice operator, and we have the precision operator $\mathcal{S}^{-1}_{\alpha,\delta,\epsilon} = \mathcal{S}_{\tilde{\alpha},\tilde{\delta},\tilde{\epsilon}}$ with $\tilde{\alpha}(t) = \alpha(t)^{-1}$ and some $\tilde{\delta}, \tilde{\epsilon} \colon [0,1] \to \mathbb{R}^{N\times L}$.

Twice the negative log likelihood may be expressed using either the covariance or the precision, i.e.

$$\begin{aligned} -2\log L(\alpha,\delta,\epsilon) &= n \log\det \hat{\mathcal{S}}_{\alpha,\delta,\epsilon} + \sum_{i=1}^n \operatorname{tr}\big\langle X_i, \hat{\mathcal{S}}^{-1}_{\alpha,\delta,\epsilon} X_i\big\rangle \\ &= -n \log\det \hat{\mathcal{S}}^{-1}_{\alpha,\delta,\epsilon} + \sum_{i=1}^n \operatorname{tr}\big\langle X_i, \hat{\mathcal{S}}^{-1}_{\alpha,\delta,\epsilon} X_i\big\rangle. \end{aligned}$$

Let $X = \{X_i^{\text{embedding}}\}_{i=1,\dots,n} \in \mathcal{H}(N,n)$ be the concatenation of the embedded observations. The variational derivatives against $f = \alpha$, δ or ϵ are given by

$$\begin{aligned} \frac{-2\partial \log L(\alpha,\delta,\epsilon)}{\partial f} &= \sum_{i=1}^n \Bigg\{ -\operatorname{tr}\bigg[\hat{\mathcal{S}}^{-1}_{\alpha,\delta,\epsilon} \frac{\partial \hat{\mathcal{S}}_{\alpha,\delta,\epsilon}}{\partial f}\bigg] + \operatorname{tr}\bigg\langle X_i, \frac{\partial \hat{\mathcal{S}}_{\alpha,\delta,\epsilon}}{\partial f} X_i \bigg\rangle \Bigg\} \\ &\approx \frac{\partial}{\partial f}\Bigg\{ -n\int_0^1 \operatorname{tr}\big[\tilde{\alpha}(s)\alpha(s)\big]\mathrm{d}s - n\int_0^1 \operatorname{tr}\big[\tilde{\alpha}(s)\delta(s)\epsilon(s)^\top\big]\mathrm{d}s \\ &\quad - n\int_0^1 \operatorname{tr}\big[\tilde{\delta}(s)\tilde{\epsilon}(s)^\top\alpha(s)\big]\mathrm{d}s - 2n\int_0^1\int_s^1 \operatorname{tr}\big[\tilde{\delta}(s)\tilde{\epsilon}(u)^\top\epsilon(u)\delta(s)^\top\big]\mathrm{d}u\,\mathrm{d}s \\ &\quad + \int_0^1 \operatorname{tr}\big[\alpha(s)X(s)X(s)^\top\big]\mathrm{d}s + 2\int_0^1\int_s^1 \operatorname{tr}\big[\delta(s)\epsilon(u)^\top X(u)X(s)^\top\big]\mathrm{d}u\,\mathrm{d}s \Bigg\} \\ &= \frac{\partial}{\partial f}\Bigg\{ -n\int_0^1 \operatorname{tr}\big[\tilde{\alpha}(s)\alpha(s)\big]\mathrm{d}s - n\int_0^1 \operatorname{tr}\big[\tilde{\alpha}(s)\delta(s)\epsilon(s)^\top\big]\mathrm{d}s \\ &\quad - n\int_0^1 \operatorname{tr}\big[\tilde{\delta}(s)\tilde{\epsilon}(s)^\top\alpha(s)\big]\mathrm{d}s - 2n\int_0^1\int_0^u \operatorname{tr}\big[\tilde{\delta}(s)\tilde{\epsilon}(u)^\top\epsilon(u)\delta(s)^\top\big]\mathrm{d}s\,\mathrm{d}u \\ &\quad + \int_0^1 \operatorname{tr}\big[\alpha(s)X(s)X(s)^\top\big]\mathrm{d}s + 2\int_0^1\int_0^u \operatorname{tr}\big[\delta(s)\epsilon(u)^\top X(u)X(s)^\top\big]\mathrm{d}s\,\mathrm{d}u \Bigg\}. \end{aligned}$$

This gives

$$\begin{aligned}
\frac{-2\partial \log L(\alpha,\delta,\epsilon)}{\partial\alpha} &= -n\tilde{\alpha} - n\frac{\tilde{\delta} * \tilde{\epsilon}^\top + \tilde{\epsilon} * \tilde{\delta}^\top}{2} + X * X^\top, \\
\frac{-2\partial \log L(\alpha,\delta,\epsilon)}{\partial\delta} &= -n\tilde{\alpha} * \epsilon - 2n\tilde{\delta} * \mathrm{B}(\tilde{\epsilon}^\top * \epsilon) + 2X * \mathrm{B}(X^\top * \epsilon), \\
\frac{-2\partial \log L(\alpha,\delta,\epsilon)}{\partial\epsilon} &= -n\tilde{\alpha} * \delta - 2n\tilde{\epsilon} * \mathrm{F}(\tilde{\delta}^\top * \delta) + 2X * \mathrm{F}(X^\top * \delta).
\end{aligned}$$

If α, δ, ϵ are piecewise linear, then derivative w.r.t. knot points may be found using triangular evaluation of the functional derivatives, i.e.

$$\begin{aligned}
\left\{\frac{-2\partial \log L(\alpha,\delta,\epsilon)}{\partial\alpha_j}\right\}_{j=1,\ldots,p} &= \mathrm{T}\left(\frac{-2\partial \log L(\alpha,\delta,\epsilon)}{\partial\alpha}\right), \\
\left\{\frac{-2\partial \log L(\alpha,\delta,\epsilon)}{\partial\delta_j}\right\}_{j=1,\ldots,p} &= \mathrm{T}\left(\frac{-2\partial \log L(\alpha,\delta,\epsilon)}{\partial\delta}\right), \\
\left\{\frac{-2\partial \log L(\alpha,\delta,\epsilon)}{\partial\epsilon_j}\right\}_{j=1,\ldots,p} &= \mathrm{T}\left(\frac{-2\partial \log L(\alpha,\delta,\epsilon)}{\partial\epsilon}\right).
\end{aligned}$$

These computations may be done rather easily using the computational engine introduced above.

□