Low-rank and sparse methods for high-dimensional approximation and model order reduction

—

Lecture 5

Tensor structure of high-dimensional equations

Introduction

In this lecture, we give the interpretation of high dimensional partial differential equations and parameter-dependent equations as operator equations in tensor spaces, and we present practical aspects for obtaining a formulation suitable for the application of tensor methods.

Ultimately, tensor-structured equations will be of the form

$$\mathbf{A}\mathbf{u} = \mathbf{b}, \quad \mathbf{u} \in \mathbb{R}^{I} = \mathbb{R}^{I_1 \times \ldots \times I_d}.$$

Outline

1. Tensor product of operators
2. Tensor structure of parameter-dependent equations
3. Tensor structure of high-dimensional PDEs

Outline

Tensor product of operators

Let $V = V^1 \otimes \ldots \otimes V^d$ and $W = W^1 \otimes \ldots \otimes W^d$ be two algebraic tensor spaces.

Let $L(V^\nu, W^\nu)$ denote the space of linear operators from V^ν to W^ν. The elementary tensor product of operators $A^{(\nu)} \in L(V^\nu, W^\nu)$, $1 \leq \nu \leq d$, denoted by

$$A = A^{(1)} \otimes \ldots \otimes A^{(d)},$$

is defined in a standard way. Then, we can define the algebraic tensor space

$$L := L(V^1, W^1) \otimes \ldots \otimes L(V^d, W^d),$$

which is the set of finite linear combinations of elementary tensors.

Tensor product of operators

For the case where

$$V = W = \mathbb{R}^I, \quad I = I_1 \times \ldots \times I_d,$$

$L(V^\nu, W^\nu)$ is identified with $\mathbb{R}^{I_\nu \times I_\nu}$ and an operator in L is identified with an element of $\mathbb{R}^{I \times I}$, such that for $u \in \mathbb{R}^I$, $Au \in \mathbb{R}^I$ is given by

$$(Au)(i) = \sum_{j \in I} A(i,j)u(j).$$

An elementary tensor $A = A^{(1)} \otimes \ldots \otimes A^{(d)}$ is such that

$$A(i,j) = A((i_1, \ldots, i_d), (j_1, \ldots, j_d)) = A^{(1)}(i_1, j_1) \ldots A^{(d)}(i_d, j_d).$$

Operators in low-rank formats

L being a tensor product of vector spaces, the ranks of tensors in L are defined in a usual way, as well as the corresponding tensor formats.

An operator A in canonical format has a representation

$$A = \sum_{k=1}^{r} A_k^{(1)} \otimes \ldots \otimes A_k^{(d)} C(k).$$

Ultimately, an operator in low-rank format has a representation of the form

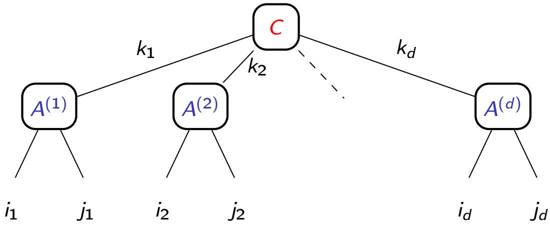
$$A = \sum_{k_1=1}^{r_1} \ldots \sum_{k_L=1}^{r_L} A_{k_{S_1}}^{(1)} \otimes \ldots \otimes A_{k_{S_d}}^{(d)} \prod_{\nu=d+1}^{M} C^{(\nu)}(k_{S_\nu}),$$

where $C^{(\nu)}$ is a tensor of order $\#S_\nu$ depending on a subset $S_\nu \subset \{1, \ldots, M\}$ of summation indices, and where the $A_{k_{S_\nu}}^{(\nu)}$ are operators in $L(V^\nu, W^\nu)$.

Operators in Tucker format

An operator A in Tucker format has a representation

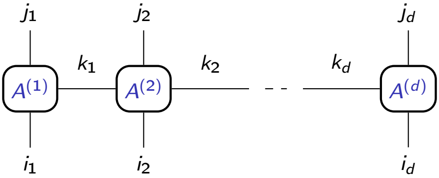
$$A = \sum_{k_1=1}^{r_1} \dots \sum_{k_d=1}^{r_d} A^{(1)}_{k_1} \otimes \dots \otimes A^{(d)}_{k_d} C(k_1, \dots, k_d).$$



Operators in Tensor Train format

An operator A in tensor train format has a representation of the form

$$A = \sum_{k_1=1}^{r_1} \dots \sum_{k_{d-1}=1}^{r_{d-1}} A^{(1)}_{1,k_1} \otimes A^{(2)}_{k_1,k_2} \otimes \dots \otimes A^{(d)}_{k_{d-1},1}.$$



Operations between tensors

For an operator A and a vector v in a low-rank format, with

$$A = \sum_{k_1=1}^{r_1} \dots \sum_{k_L=1}^{r_L} \left(\bigotimes_{\nu=1}^{d} A^{(\nu)}_{k_{S_\nu}} \right) \prod_{\nu=d+1}^{M} C^{(\nu)}(k_{S_\nu}),$$

$$v = \sum_{k_1=1}^{r'_1} \dots \sum_{k_{L'}=1}^{r'_{L'}} \left(\bigotimes_{\nu=1}^{d} v^{(\nu)}_{k_{S'_\nu}} \right) \prod_{\nu=d+1}^{M'} D^{(\nu)}(k_{S'_\nu}),$$

the product Av is a tensor such that

$$Av = \sum_{k_1=1}^{r_1} \dots \sum_{k_L=1}^{r_L} \sum_{k'_1=1}^{r'_1} \dots \sum_{k_{L'}=1}^{r'_{L'}} \left(\bigotimes_{\nu=1}^{d} \left(A^{(\nu)}_{k_{S_\nu}} v^{(\nu)}_{k_{S'_\nu}} \right) \right) \prod_{\nu=d+1}^{M} C^{(\nu)}(k_{S_\nu}) \prod_{\nu=d+1}^{M'} D^{(\nu)}(k_{S'_\nu})$$

Operations between tensors

For two tensors u and v in a Hilbert tensor space V, with

$$u = \sum_{k_1=1}^{r_1} \dots \sum_{k_L=1}^{r_L} \left(\bigotimes_{\nu=1}^{d} u^{(\nu)}_{k_{S_\nu}} \right) \prod_{\nu=d+1}^{M} C^{(\nu)}(k_{S_\nu}),$$

$$v = \sum_{k_1=1}^{r'_1} \dots \sum_{k_{L'}=1}^{r'_{L'}} \left(\bigotimes_{\nu=1}^{d} v^{(\nu)}_{k_{S'_\nu}} \right) \prod_{\nu=d+1}^{M'} D^{(\nu)}(k_{S'_\nu}),$$

the inner product of u and v is such that

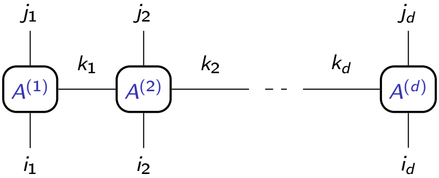
$$(u, v) = \sum_{k_1=1}^{r_1} \dots \sum_{k_L=1}^{r_L} \sum_{k'_1=1}^{r'_1} \dots \sum_{k_{L'}=1}^{r'_{L'}} \left(\prod_{\nu=1}^{d} \left(u^{(\nu)}_{k_{S_\nu}}, v^{(\nu)}_{k_{S'_\nu}} \right)_\nu \right) \prod_{\nu=d+1}^{M} C^{(\nu)}(k_{S_\nu}) \prod_{\nu=d+1}^{M'} D^{(\nu)}(k_{S'_\nu}).$$

Graphical representation of operations for tree-based or tensor networks formats

Operations between tensors in tree-based tensor formats or more general tensor networks formats have a simple graphical representation. For example, consider an operator A and a vector v in tensor train format

$$A(i,j) = \sum_{k_1=1}^{r_1} \dots \sum_{k_{d-1}=1}^{r_{d-1}} A^{(1)}(i_1, j_1, k_1) A^{(2)}(i_2, j_2, k_1, k_2) \dots A^{(d)}(i_d, j_d, k_{d-1})$$

$$v(j) = \sum_{k_1=1}^{r'_1} \dots \sum_{k_{d-1}=1}^{r'_{d-1}} v^{(1)}(j_1, k_1) v^{(2)}(j_2, k_1, k_2) \dots v^{(d)}(j_d, k_{d-1}).$$



Parameter-dependent equations

Let us consider a parameter-dependent operator equation

$$A(\xi)u(\xi) = b(\xi), \tag{1}$$

where $\xi = (\xi_1, \ldots, \xi_s)$ are parameters or random variables taking values in Ξ,

$$A(\xi) : \mathcal{V} \to \mathcal{W}$$

is a parameter-dependent linear operator, and

$$b(\xi) \in \mathcal{W}$$

is a parameter-dependent vector.

Affine representations

We here assume that $A(\xi)$ and $b(\xi)$ admit so-called affine representations

$$A(\xi) = \sum_{i=1}^{L} \lambda_i(\xi) A_i, \quad b(\xi) = \sum_{i=1}^{R} \eta_i(\xi) b_i, \tag{2}$$

with $A_i : \mathcal{V} \to \mathcal{W}$ and $b_i \in \mathcal{W}$.

Example 1 (Diffusion-reaction equation)

The problem

$$-\lambda_1(\xi)\Delta u + \lambda_2(\xi) u = \eta_1(\xi) b_1 \quad \text{on } D, \quad u = 0 \quad \text{on } \partial D,$$

can be written in the form $A(\xi)u(\xi) = b(\xi)$, where $A(\xi)$ has an affine representation with $L = 2$, $A_1 v = -\Delta v$ and $A_2 v = v$, and where $b(\xi)$ has an affine representation with $R = 1$.

Remark.

Some problems have operators and right-hand side directly given in the form (2). If this is not the case (or if R and L are high), a preliminary approximation step is required (e.g. using interpolation).

Affine representations

Example 2 (Diffusion equation with random diffusion coefficient)

The problem

$$-\nabla \cdot (k(\cdot,\omega)\nabla u) = b \quad \text{on} \quad D, \quad u = 0 \quad \text{on} \quad \partial D,$$

where $k(x,\omega)$ is a second random field with a decomposition

$$k(x,\omega) = \sum_{i=1}^{L} k_i(x)\xi_i(\omega), \quad L \in \mathbb{N} \cup \{+\infty\},$$

can be written in the form $A(\xi)u(\xi) = b$, where

$$A(\xi) = \sum_{i=1}^{L} A_i \xi_i \quad \text{with} \quad A_i v = -\nabla \cdot (k_i \nabla v).$$

Affine representations

Example 3 (Diffusion equation on a random domain)

Consider the problem

$$-\Delta U(x,\xi) = g(x) \quad \text{for} \quad x \in D(\xi), \qquad\qquad U(x,\xi) = 0 \quad \text{for} \quad x \in \partial D(\xi).$$

Assume that there exists a diffeomorphism $\phi(\cdot;\xi) : D_0 \to D(\xi)$ from a deterministic domain D_0 to $D(\xi)$. By using the change of variable

$$u(x_0,\xi) = U(\phi(x_0,\xi),\xi), \; x_0 \in D_0,$$

the problem can be interpreted as a diffusion equation on a deterministic domain but with random diffusion coefficient and source term:

$$-\nabla \cdot (K(\cdot,\xi)\nabla u) = g_0(\cdot,\xi),$$

with

$$K(x_0,\xi) = \nabla\phi(x_0,\xi)\nabla\phi(x_0,\xi)^T|\det(\nabla\phi(x_0,\xi))|$$
$$g_0(x_0,\xi) = g(\phi(x_0,\xi))|\det(\nabla\phi(x_0,\xi))|.$$

Apart from simple transformations ϕ (e.g. affine), approximations of K and g_0 are required to obtain affine representations of the operator and right-hand side.

Parameter-dependent equations

For simplicity, let us assume that $\mathcal{V}$ and $\mathcal{W}$ are N-dimensional spaces and identify the equation

$$A(\xi)u(\xi) = b(\xi),$$

with a linear system of equations

$$\mathbf{A}(\xi)\mathbf{u}(\xi) = \mathbf{b}(\xi),$$

with

$$\mathbf{A}(\xi) \in \mathbb{R}^{N\times N}, \quad \mathbf{u}(\xi) \in \mathbb{R}^{N}, \quad \mathbf{b}(\xi) \in \mathbb{R}^{N}.$$

Example 4 (Diffusion-reaction equation)

In example 1, consider that $\mathcal{V} = \mathcal{W}$ is an approximation space in $H^1_0(D)$ (e.g. a finite element space) with basis $\{\varphi_i\}_{i=1}^N$, and let $u \in \mathcal{V}$ be the standard Galerkin approximation of the solution of the PDE.

Then $\mathbf{u}(\xi)$ are the coefficients of u on the basis of $\mathcal{V}$, and $\mathbf{A}(\xi)$ and $\mathbf{b}(\xi)$ admit affine representations

$$\mathbf{A}(\xi) = \mathbf{A}_1\lambda_1(\xi) + \mathbf{A}_2\lambda_2(\xi) \quad \text{and} \quad \mathbf{b}(\xi) = \mathbf{b}_1\eta_1(\xi),$$

with

$$\mathbf{A}_1(i,j) = \int_D \nabla\varphi_i \cdot \nabla\varphi_j, \quad \mathbf{A}_2(i,j) = \int_D \varphi_i\varphi_j, \quad \mathbf{b}_1(i) = \int_D \varphi_i b_1.$$

Parameter-dependent equations for a finite training set

We also assume that we are interested in evaluating the solution $\mathbf{u}(\xi)$ at a finite set of values $\{\xi^k\}_{k\in K}$ of ξ (a training set) , such that

$$\mathbf{A}(\xi^k)\mathbf{u}(\xi^k) = \mathbf{b}(\xi^k), \quad \forall k \in K. \tag{3}$$

The set of vectors $\{\mathbf{u}(\xi^k)\}_{k\in K}$ and $\{\mathbf{b}(\xi^k)\}_{k\in K}$, as elements of $(\mathbb{R}^N)^K$, can be identified with order-two tensors

$$\mathbf{u} \in \mathbb{R}^N \otimes \mathbb{R}^K \quad \text{and} \quad \mathbf{b} \in \mathbb{R}^N \otimes \mathbb{R}^K.$$

The set of matrices $\{\mathbf{A}(\xi^k)\}_{k\in K}$, considered as a linear operator from $\mathbb{R}^N \otimes \mathbb{R}^K$ and $\mathbb{R}^N \otimes \mathbb{R}^K$, can be identified with a tensor

$$\mathbf{A} \in \mathbb{R}^{N\times N} \otimes \mathbb{R}^{K\times K}.$$

Finally, the set of equations (3) can be identified with a operator equation

$$\mathbf{A}\mathbf{u} = \mathbf{b}.$$

Parameter-dependent equations for a finite training set

The affine representations of parameter-dependent operator $A(\xi)$ yields a low-rank representation for the tensor **A** in the form

$$\mathbf{A} = \sum_{i=1}^{L} \mathbf{A}_i \otimes \mathbf{\Lambda}_i, \quad \text{with } \mathbf{\Lambda}_i = \text{diag}(\boldsymbol{\lambda}_i), \quad \boldsymbol{\lambda}_i = (\lambda_i(\xi^k))_{k \in K}.$$

Also, the affine representation of parameter-dependent vector $b(\xi)$ yields a low-rank representation for the tensor **b** in the form

$$\mathbf{b} = \sum_{i=1}^{R} \mathbf{b}_i \otimes \boldsymbol{\eta}_i, \quad \boldsymbol{\eta}_i = (\eta_i(\xi^k))_{k \in K}.$$

Parameter-dependent equations for a tensorized training set

Let us assume that $\xi = (\xi_1, \ldots, \xi_s)$ is a vector of parameters taking values in a product set $\Xi = \Xi_1 \times \ldots \times \Xi_s$.

Let $\{\xi_\nu^{k_\nu}\}_{k_\nu \in K_\nu}$ be a grid in Ξ_ν, and let us consider for the training set the tensorized grid

$$\{\xi^k = (\xi_1^{k_1}, \ldots, \xi_s^{k_s})\}_{k \in K}, \quad K = K_1 \times \ldots \times K_d.$$

A vector $\mathbf{a} \in \mathbb{R}^K$ is then identified with a tensor in $\mathbb{R}^{K_1} \otimes \ldots \otimes \mathbb{R}^{K_s}$.

Then the tensor $\mathbf{u} \in \mathbb{R}^N \otimes \mathbb{R}^K$ can be identified with a higher-order tensor

$$\mathbf{u} \in \mathbb{R}^N \otimes \mathbb{R}^{K_1} \otimes \ldots \otimes \mathbb{R}^{K_s}.$$

Parameter-dependent equations for a tensorized training set

Let us assume that $\mathbf{A}(\xi)$ and $\mathbf{b}(\xi)$ admit affine representations

$$\mathbf{A}(\xi) = \sum_{i=1}^{L} \mathbf{A}_i \lambda_i(\xi) \quad \text{and} \quad \mathbf{b}(\xi) = \sum_{i=1}^{R} \mathbf{b}_i \eta_i(\xi),$$

with rank-one functions

$$\lambda_i(\xi) = \lambda_i^{(1)}(\xi_1)\dots\lambda_i^{(s)}(\xi_s) \quad \text{and} \quad \eta_i(\xi) = \eta_i^{(1)}(\xi_1)\dots\eta_i^{(s)}(\xi_s).$$

The set of evaluations of a rank one function $a(\xi) = a^{(1)}(\xi_1)\dots a^{(s)}(\xi_s)$ on the tensorized grid is identified with a rank-one tensor

$$\mathbf{a} = (a(\xi^k))_{k\in K} = \mathbf{a}^{(1)} \otimes \dots \otimes \mathbf{a}^{(s)}, \quad \mathbf{a}^{(\nu)} = (a^{(\nu)}(\xi_\nu^{k_\nu}))_{k_\nu \in K_\nu}.$$

Parameter-dependent equations for a tensorized training set

Then the problem can be interpreted as a tensor-structured equation

$$\mathbf{A}\mathbf{u} = \mathbf{b}$$

with $\mathbf{u} \in \mathbb{R}^N \otimes \mathbb{R}^{K_1} \otimes \ldots \otimes \mathbb{R}^{K_d}$, and where $\mathbf{A}$ and $\mathbf{b}$ admit low-rank representations (in canonical format)

$$\mathbf{b} = \sum_{i=1}^{L} \mathbf{b}_i \otimes \boldsymbol{\eta}_i^{(1)} \otimes \ldots \otimes \boldsymbol{\eta}_i^{(s)}$$

and

$$\mathbf{A} = \sum_{i=1}^{R} \mathbf{A}_i \otimes \boldsymbol{\Lambda}_i^{(1)} \otimes \ldots \otimes \boldsymbol{\Lambda}_i^{(s)}, \quad \boldsymbol{\Lambda}_i^{(\nu)} = \mathrm{diag}(\boldsymbol{\lambda}_i^{(\nu)}).$$

Galerkin methods

Consider the problem of finding $\mathbf{u} : \Xi \to \mathbb{R}^N$ such that

$$\mathbf{A}(\xi)\mathbf{u}(\xi) = \mathbf{b}(\xi) \quad \text{for all } \xi \text{ in } \Xi,$$

where Ξ is equipped with a measure P_ξ.

We introduce a finite dimensional space $\mathcal{S}$ of functions defined on Ξ.

The Galerkin approximation $\tilde{\mathbf{u}}$ of $\mathbf{u}$ in the space $\mathbb{R}^N \otimes \mathcal{S}$, also denoted $\mathbf{u}$, is defined by

$$\mathbf{a}(\mathbf{u}, \mathbf{v}) = \boldsymbol{\ell}(\mathbf{v}), \quad \forall \mathbf{v} \in \mathbb{R}^N \otimes \mathcal{S},$$

where $\mathbf{a}$ is a linear form defined by

$$\mathbf{a}(\mathbf{u}, \mathbf{v}) = \int_\Xi \langle \mathbf{A}(y)\mathbf{u}(y), \mathbf{v}(y) \rangle P_\xi(dy)$$

and $\boldsymbol{\ell}$ is a linear form defined by

$$\boldsymbol{\ell}(\mathbf{v}) = \int_\Xi \langle \mathbf{b}(y), \mathbf{v}(y) \rangle P_\xi(dy).$$

Galerkin methods

Let $\{\psi_k(\xi)\}_{k\in K}$ be a basis of $\mathcal{S}$. The tensor $\mathbf{u} = \sum_{k\in K} \mathbf{u}_k \otimes \psi_k$ in $\mathbb{R}^N \otimes \mathcal{S}$ can be identified with a tensor $\mathbf{U} = (\mathbf{u}_k)_{k\in K} \in \mathbb{R}^N \otimes \mathbb{R}^K$.

Finally, the problem defining the Galerkin approximation is identified with a tensor structured equation on $\mathbf{U} \in \mathbb{R}^N \otimes \mathbb{R}^K$,

$$\mathbf{A}\mathbf{U} = \mathbf{b},$$

where $\mathbf{A} \in \mathbb{R}^{N\times N} \otimes \mathbb{R}^{K\times K}$ and $\mathbf{b} \in \mathbb{R}^N \otimes \mathbb{R}^K$ have low-rank representations

$$\mathbf{A} = \sum_{i=1}^{R} \mathbf{A}_i \otimes \mathbf{\Lambda}_i \quad \text{and} \quad \mathbf{b} = \sum_{i=1}^{L} \mathbf{b}_i \otimes \boldsymbol{\eta}_i,$$

with $\mathbf{\Lambda}_i \in \mathbb{R}^{K\times K}$ such that

$$\mathbf{\Lambda}_i(k,l) = \int_\Xi \psi_k(y)\psi_l(y)\lambda_i(x)P_\xi(dy),$$

and $\boldsymbol{\eta}_i \in \mathbb{R}^K$ such that

$$\boldsymbol{\eta}_i(k) = \int_\Xi \eta_i(y)\psi_k(y)P_\xi(dy).$$

Outline

High-dimensional partial differential equations

Let $\mathcal{X}$ in $\mathbb{R}^d$ be a product domain of $\mathbb{R}^d$, with

$$\mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_d.$$

Let us consider the problem of finding a multivariate function

$$u(x_1, \ldots, x_d)$$

which satisfies suitable boundary conditions on $\partial\mathcal{X}$ and a partial differential equation

$$A(u) = b \quad \text{on} \quad \mathcal{X},$$

where b is a given multivariate function and A is an operator such that $A(u)$ depends on the partial derivatives

$$D^\alpha u = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}} u,$$

where $|\alpha| := \|\alpha\|_1$ is the length of the multi-index $\alpha \in \mathbb{N}^d$.

Example 5 (Laplace operator)

$$\Delta u = \frac{\partial^2}{\partial x_1^2} u + \ldots + \frac{\partial^2}{\partial x_d^2} u = D^{(2,0\ldots,0)} u + \ldots + D^{(0,\ldots,0,2)} u$$

High-dimensional partial differential equations

Remark. Non product domains

A partial differential equation $A(u) = b$ defined on domain $\mathcal{X}$ which is not a product domain can be transformed into a partial differential equation on a product domain in two different ways:

- by introducing a bijection $\phi : \tilde{\mathcal{X}} \to \mathcal{X}$ from a product domain $\tilde{\mathcal{X}}$ to $\mathcal{X}$ and by using a change of variable $u(x) = \tilde{u}(\phi(x))$. The map ϕ should be sufficiently smooth (possibly piecewise smooth).
- by embedding the domain $\mathcal{X}$ into a fictitious product domain $\tilde{\mathcal{X}}$, and using consistant reformulations of the problem on the fictitious domain.

Tensor structure of differential operators

Assume that the problem admits a unique solution u in a space $\overline{V}^{\|\cdot\|}$ where $V = V^1 \otimes \ldots \otimes V^d$ is the tensor product of spaces V^ν of functions defined on $\mathcal{X}_\nu$.

For an elementary tensor

$$v(x) = v^{(1)}(x_1) \ldots v^{(d)}(x_d),$$

and for $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$, the differential operator D^α is such that

$$D^\alpha v(x) = D^{\alpha_1} v^{(1)}(x_1) \ldots D^{\alpha_d} v^{(d)}(x_d).$$

Then D^α can be interpreted as an elementary operator on the tensor space V, with

$$D^\alpha = D^{\alpha_1} \otimes \ldots \otimes D^{\alpha_d}.$$

Differential operators in low-rank tensor formats

A linear partial differential operator of the form

$$A = \sum_{\alpha} a_\alpha D^\alpha,$$

where $a_\alpha \in \mathbb{R}$, can then be identified with an operator on V with admits a representation in canonical format

$$A = \sum_{\alpha} a_\alpha D^{\alpha_1} \otimes \ldots \otimes D^{\alpha_d}.$$

Example 6 (Laplace operator)

The Laplace operator is identified with a tensor with canonical rank d

$$\Delta = D^2 \otimes I \otimes \ldots \otimes I + \ldots + I \otimes \ldots \otimes I \otimes D^2,$$

Differential operators in low-rank tensor formats

Differential operators may have representations with reduced complexity in other tensor formats.

Example 7 (Laplace operator in tensor train format)

The Laplace operator admits a representation in tensor train format with TT-rank $(2, \dots, 2)$

$$\Delta = \sum_{k_1=1}^{2} \dots \sum_{k_{d-1}=1}^{2} B_{2,k_1} \otimes B_{k_1,k_2} \dots \otimes B_{k_{d-1},1}$$

where

$$B_{1,1} = B_{2,2} = I, \quad B_{1,2} = 0, \quad B(2,1) = D^2.$$

This can be represented in a more convenient block form where each block represents a collection of operators $\{B_{k_1,k_2}\}$

$$\Delta = \begin{pmatrix} D^2 & I \end{pmatrix} \bowtie \begin{pmatrix} I & 0 \\ D^2 & I \end{pmatrix} \bowtie \dots \bowtie \begin{pmatrix} I & 0 \\ D^2 & I \end{pmatrix} \bowtie \begin{pmatrix} I \\ D^2 \end{pmatrix}.$$

where

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \bowtie \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11} \otimes B_{11} + A_{12} \otimes B_{21} & A_{11} \otimes B_{12} + A_{12} \otimes B_{22} \\ A_{21} \otimes B_{11} + A_{22} \otimes B_{21} & A_{21} \otimes B_{12} + A_{22} \otimes B_{22} \end{pmatrix}$$

Differential operators in low-rank tensor formats

Example 8 (Laplace operator in tree-based tensor format)

The Laplace operator admits a representation in Tucker format with rank $(2, \dots, 2)$, such that

$$\Delta = \sum_{k_1=1}^{2} \dots \sum_{k_d=1}^{2} B_{k_1} \otimes \dots \otimes B_{k_d} C(k_1, \dots, k_d),$$

where $B_1 = I$ and $B_2 = D^2$, and where the tensor $C \in \mathbb{R}^2 \otimes \dots \otimes \mathbb{R}^2$ has a representation in canonical format

$$C = e_2 \otimes e_1 \otimes \dots \otimes e_1 + \dots + e_1 \otimes e_2 \otimes \dots \otimes e_2, \quad e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The tensor C has a exact representation in tree-based format $\mathcal{T}_r^T$ with T-rank $r = (2, \dots, 2)$ whatever the tree T.

Differential operators in low-rank tensor formats

Example 9 (Representation of Laplace-like operators in tensor train format)

An operator A of the form

$$A = M_1 \otimes R_2 \otimes R_3 \otimes \dots R_d + L_1 \otimes M_2 \otimes R_3 \otimes \dots \otimes R_d + \dots \\ + L_1 \otimes \dots \otimes L_{d-2} \otimes M_{d-1} \otimes R_d + L_1 \otimes \dots \otimes L_{d-1} \otimes M_d$$

admits a representation in tensor train format with TT rank $(2, \dots, 2)$

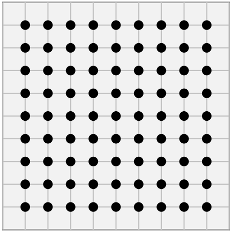
$$A = \begin{pmatrix} L_1 & M_1 \end{pmatrix} \bowtie \begin{pmatrix} L_2 & M_2 \\ & R_2 \end{pmatrix} \bowtie \dots \bowtie \begin{pmatrix} L_{d-1} & M_{d-1} \\ & R_{d-1} \end{pmatrix} \bowtie \begin{pmatrix} M_d \\ R_d \end{pmatrix}.$$

Finite difference schemes on tensorized grids

Let consider uniform uni-dimensional grids $\Gamma^{\nu}_{I_\nu} = \{x_\nu^{i_\nu}\}_{i_\nu \in I_\nu}$ in $\mathcal{X}_\nu$.

Let $I = I_1 \times \ldots \times I_d$ and let Γ_I be the tensorized grid on $\mathcal{X}$ defined by

$$\Gamma_I = \Gamma^1_{I_1} \times \ldots \times \Gamma^d_{I_d} = \{x^i = (x_1^{i_1}, \ldots, x_1^{i_d}) : i \in I\} \ .$$



Finite difference schemes on tensorized grids

For a function $v(x_\nu)$ on $\mathcal{X}_\nu$ with values $\mathbf{v} = (v(x_\nu^{i_\nu}))_{i_\nu \in I_\nu}$ on the grid $\Gamma^\nu_{I_\nu}$, we define a finite difference operator

$$\mathbf{D}^k_\nu : \mathbb{R}^{I_\nu \times I_\nu}$$

associated with D^k, such that $\mathbf{D}^k\mathbf{v}$ provides a finite difference approximation of $D^k v(x_\nu)$ on the grid $\Gamma^\nu_{I_\nu}$.

Example 10

For $\mathcal{X}_\nu = (0,1)$, and a uniform grid $\Gamma_\nu = \{ih\}_{i=1}^n$ of step size $h = (n+1)^{-1}$, a standard difference operators for D^2 (for Dirichlet boundary conditions) is given by

$$\mathbf{D}^2_\nu = -h^{-2} \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}$$

Then, for $\alpha \in \mathbb{N}^d$,

$$\mathbf{D}^\alpha = \mathbf{D}_1^{\alpha_1} \otimes \ldots \otimes \mathbf{D}_d^{\alpha_d} \in \mathbb{R}^{I_1 \times I_1} \otimes \ldots \otimes \mathbb{R}^{I_d \times I_d}$$

is a finite difference operator associated with the differential operator D^α on the tensorized grid Γ_I.

Finite difference schemes on tensorized grids

A finite difference scheme for the linear partial differential equation

$$Au = b,$$

with

$$A = \sum_{\alpha} a_\alpha D^\alpha,$$

yields a tensor structured equation

$$\mathbf{A}\mathbf{u} = \mathbf{b},$$

where the entries of the tensor

$$\mathbf{u} \in \mathbb{R}^I = \mathbb{R}^{I_1 \times \ldots \times I_d}$$

provide approximations of the values of the solution $u(x)$ on the tensor product grid Γ_I, where

$$\mathbf{b} = (b(x^i))_{i \in I} \in \mathbb{R}^I$$

is the vector of evaluations of the function $b(x)$ on the tensor product grid Γ_I, and where **A** is an operator with the following representation in canonical format

$$\mathbf{A} = \sum_{\alpha} a_\alpha \mathbf{D}_1^{\alpha_1} \otimes \ldots \otimes \mathbf{D}_d^{\alpha_d}.$$

Finite difference schemes on tensorized grids

Example 11 (Discrete Laplace operator in low-rank formats)

We consider uniform grids $\Gamma^{\nu}_{I_\nu}$ with n points and step size h. The Discrete Laplace operator $\mathbf{\Delta}$ admits a representation in canonical format

$$\mathbf{\Delta} = \mathbf{B} \otimes \mathbf{I} \otimes \ldots \otimes \mathbf{I} + \ldots + \mathbf{I} \otimes \ldots \otimes \mathbf{I} \otimes \mathbf{B},$$

with $\mathbf{B} = h^{-2}\,\mathrm{diag}(-1, 2, -1)$ for Dirichlet boundary conditions.
It also admits a representation in tensor train format with TT-rank $(2, \ldots, 2)$

$$\mathbf{\Delta} = \sum_{k_1=1}^{2} \ldots \sum_{k_{d-1}=1}^{2} \mathbf{B}_{2,k_1} \otimes \mathbf{B}_{k_1,k_2} \ldots \otimes \mathbf{B}_{k_{d-1},1}$$

with $\mathbf{B}_{1,1} = \mathbf{B}_{2,2} = \mathbf{I}$, $\quad \mathbf{B}_{1,2} = 0$, $\quad \mathbf{B}_{1,2} = \mathbf{B}$, or using a block notation,

$$\mathbf{\Delta} = \begin{pmatrix} \mathbf{B} & \mathbf{I} \end{pmatrix} \bowtie \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{B} & \mathbf{I} \end{pmatrix} \bowtie \ldots \bowtie \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{B} & \mathbf{I} \end{pmatrix} \bowtie \begin{pmatrix} \mathbf{I} \\ \mathbf{B} \end{pmatrix}.$$

Finite difference schemes on tensorized grids

Example 12 (Diffusion reaction equation)

Consider the equation

$$-\Delta u + \beta u = b,$$

with $b(x) = b^{(1)}(x_1) \dots b^{(d)}(x_d)$, and homogeneous Dirichlet boundary conditions.

We consider uniform grids $\Gamma^{\nu}_{I_\nu}$ with n points and step size h.

A standard (centered) finite difference scheme yields a system

$$\mathbf{A}\mathbf{u} = \mathbf{b}$$

with

$$\mathbf{b} = \mathbf{b}^{(1)} \otimes \dots \otimes \mathbf{b}^{(d)},$$

where $\mathbf{b}^{(\nu)} \in \mathbb{R}^n$ is the vector of evaluations of function $b^{(\nu)}$ on the grid $\Gamma^{\nu}_{I_\nu}$,

$$\mathbf{A} = \boldsymbol{\Delta} + \beta \mathbf{I} \otimes \dots \otimes \mathbf{I}.$$

Some details about the functional framework

Under standard assumptions, the problem is proved to be well-posed, with a solution u is the Sobolev space

$$H^k(\mathcal{X})$$

of functions u with weak partial derivatives $D^\alpha u$ in $L^2(\mathcal{X})$, for $|\alpha| \leq k$.

The space $H^k(\mathcal{X})$, equipped with the norm

$$\|u\|_{H_k}^2 = \sum_{|\alpha|\leq k} \|D^\alpha u\|_{L^2}^2,$$

is a Hilbert space which can be identified with the completion of the algebraic tensor space $H^k(\mathcal{X}_1) \otimes \ldots \otimes H^k(\mathcal{X}_d)$ with respect to the norm $\|u\|_{H_k}$, that means

$$H^k(\mathcal{X}) = \overline{H^k(\mathcal{X}_1) \otimes \ldots \otimes H^k(\mathcal{X}_d)}^{\|\cdot\|_{H^k}}.$$

Some details about the functional framework

The canonical norm on the algebraic tensor product space $H^k(\mathcal{X}_1) \otimes \ldots \otimes H^k(\mathcal{X}_d)$, which is induced by the norms on the spaces $H^k(\mathcal{X}_\nu)$, corresponds to the H^k_{mix} norm defined by

$$\|v\|^2_{H^k_{mix}} = \sum_{\|\alpha\|_\infty \leq k} \|D^\alpha v\|^2_{L^2},$$

and such that for $v(x) = v^{(1)}(x_1) \ldots v^{(d)}(x_d)$,

$$\|v\|_{H^k_{mix}} = \prod_{\nu=1}^{d} \|v^{(\nu)}\|_{H^k}.$$

Noting that $\|v\|_{H^k} \leq \|v\|_{H^k_{mix}}$, we have that the tensor space

$$H^k_{mix}(\mathcal{X}) = \overline{H^k(\mathcal{X}_1) \otimes \ldots \otimes H^k(\mathcal{X}_d)}^{\|\cdot\|_{H^k_{mix}}}$$

is such that

$$H^k_{mix}(\mathcal{X}) \subset H^k(\mathcal{X}),$$

with strict inclusion. The spaces H^k_{mix} with mixed Sobolev regularity play an important role in the analysis of approximation methods in high-dimension (sparse grids, low-rank approximations).

Galerkin methods

Assume that the problem admits a weak solution $u \in \mathcal{V}$, where $\mathcal{V}$ is a Hilbert space of functions in $H^k(\mathcal{X})$, such that

$$a(u, v) = \ell(v) \quad \forall v \in \mathcal{V},$$

where $a = \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ is a bilinear form and $\ell = \mathcal{V} \to \mathbb{R}$ a linear form.

Let $V = V^1 \otimes \ldots \otimes V^d$ be an approximation space in $\mathcal{V}$, with $V^\nu \subset H^k(\mathcal{X}_\nu)$.

A standard Galerkin projection method defines an approximation $\tilde{u}$ of u in V by

$$a(\tilde{u}, v) = \ell(v) \quad \forall v \in V,$$

Letting $\{\phi_i = \psi^{(1)}_{i_1} \otimes \ldots \otimes \psi^{(d)}_{i_d}\}_{i \in I}$ be a tensor product basis of V, the Galerkin projection is defined by the equation

$$\mathbf{A}\mathbf{u} = \mathbf{b},$$

where the tensor $\mathbf{u} \in \mathbb{R}^I$ is the set of coefficients of $\tilde{u}$ on the tensor product basis, and where $\mathbf{A} \in \mathbb{R}^{I \times I}$ and $\mathbf{b} \in \mathbb{R}^I$ are defined by

$$\mathbf{A}(i, j) = a(\psi_j, \psi_i), \quad \mathbf{b}(i) = \ell(\psi_i).$$

Galerkin methods

The tensor structure of the operator $\mathbf{A} \in \mathbb{R}^{I_1 \times I_1} \otimes \ldots \mathbb{R}^{I_d \times I_d}$ and of the right-hand side $\mathbf{b} \in \mathbb{R}^{I_1} \otimes \ldots \otimes \mathbb{R}^{I_d}$ can be exhibited by considering the action of a and ℓ on rank-one functions.

Assuming that

$$a(u^{(1)} \otimes \ldots \otimes u^{(d)}, v^{(1)} \otimes \ldots \otimes v^{(d)}) = \sum_{k=1}^{L} a_k^{(1)}(u^{(1)}, v^{(1)}) \ldots a_k^{(d)}(u^{(d)}, v^{(d)}),$$

the operator $\mathbf{A}$ has the following representation in canonical format:

$$\mathbf{A} = \sum_{k=1}^{L} \mathbf{A}_k^{(1)} \otimes \ldots \otimes \mathbf{A}_k^{(d)},$$

where $\mathbf{A}_k^{(\nu)} \in \mathbb{R}^{I_\nu \times I_\nu}$ is the operator associated with the bilinear form $a_k^{(\nu)}$, such that

$$\mathbf{A}_k^{(\nu)}(i_\nu, j_\nu) = a_k^{(\nu)}(\psi_{j_\nu}^{(\nu)}, \psi_{i_\nu}^{(\nu)}).$$

Galerkin methods

Assuming that

$$\ell(v^{(1)} \otimes \ldots \otimes v^{(d)}) = \sum_{k=1}^{R} \ell_k^{(1)}(v^{(1)}) \ldots \ell_k^{(d)}(v^{(d)}),$$

the right-hand side $\mathbf{b}$ has the following representation in canonical format:

$$\mathbf{b} = \sum_{k=1}^{R} \mathbf{b}_k^{(1)} \otimes \ldots \otimes \mathbf{b}_k^{(d)},$$

where $\mathbf{b}_k^{(\nu)} \in \mathbb{R}^{I_\nu}$ is the vector associated with the linear form $\ell_k^{(\nu)}$, such that

$$\mathbf{b}_k^{(\nu)}(i_\nu) = \ell_k^{(\nu)}(\psi_{i_\nu}^{(\nu)}).$$

Galerkin methods

Example 13 (Diffusion reaction equation)

Consider the equation

$$-\Delta u + \beta u = b \quad \text{on } \mathcal{X},$$

with $b(x) = b^{(1)}(x_1)\dots b^{(d)}(x_d)$, and homogeneous Dirichlet boundary conditions. The problem admits a weak solution $u \in \mathcal{V} = H_0^1(\mathcal{X})$ such that $a(u,v) = \ell(v)\ \forall v \in \mathcal{V}$, with

$$a(u,v) = \int_{\mathcal{X}} (\nabla u \cdot \nabla v + \beta uv), \quad \ell(v) = \int_{\mathcal{X}} bv.$$

We have

$$a(\bigotimes_\nu u^{(\nu)}, \bigotimes_\nu v^{(\nu)}) = \sum_{\nu=1}^d \int_{\mathcal{X}_\nu} \partial_{x_\nu} u^{(\nu)} \partial_{x_\nu} v^{(\nu)} \int_{\times_{\eta\neq\nu}\mathcal{X}_\eta} u^{(\eta)} v^{(\eta)} + \beta \prod_{\nu=1}^d \int_{\mathcal{X}_\nu} u^{(\nu)} v^{(\nu)},$$

$$\ell(\bigotimes_\nu v^{(\nu)}) = \prod_{\nu=1}^d \int_{\mathcal{X}_\nu} b^{(\nu)} v^{(\nu)},$$

which yields

$$\mathbf{A} = \mathbf{B}^{(1)} \otimes \mathbf{M}^{(2)} \otimes \dots \otimes \mathbf{M}^{(d)} + \dots + \mathbf{M}^{(1)} \otimes \dots \otimes \mathbf{M}^{(d-1)} \otimes \mathbf{B}^{(d)} + \beta \mathbf{M}^{(1)} \otimes \dots \otimes \mathbf{M}^{(d)},$$

with $\mathbf{B}^\nu(i_\nu, j_\nu) = \int_{\mathcal{X}_\nu} \partial_{x_\nu} \psi_{i_\nu}^{(\nu)} \partial_{x_\nu} \psi_{j_\nu}^{(\nu)}$ and $\mathbf{M}^{(\nu)}(i_\nu, j_\nu) = \int_{\mathcal{X}_\nu} \psi_{i_\nu}^{(\nu)} \psi_{j_\nu}^{(\nu)}$, and

$$\mathbf{b} = \mathbf{b}^{(1)} \otimes \dots \otimes \mathbf{b}^{(d)},$$

with $\mathbf{b}^{(\nu)}(i_\nu) = \int_{\mathcal{X}_\nu} b^{(\nu)} \psi_{i_\nu}^{(\nu)}$.