Numerical methods for PDEs, IESC, Cargèse, 2016

Tensor numerical methods for high-dimensional problems

—

Part 3

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

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

Outline

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

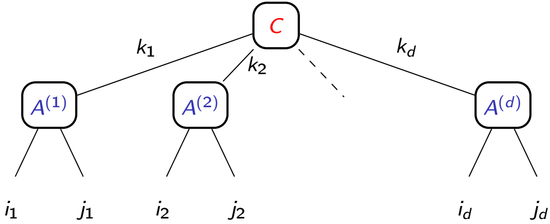
$$A = \sum_{k_1=1}^{r_1} \ldots \sum_{k_L=1}^{r_L} A_{k_{S_1}}^{(1)} \otimes \ldots \otimes A_{k_{S_d}}^{(d)} \prod_{\nu=d+1}^{M} C^{(\nu)}(k_{S_\nu}),$$

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

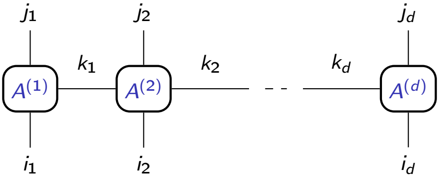
$$A = \sum_{k_1=1}^{r_1} \dots \sum_{k_d=1}^{r_d} A^{(1)}_{k_1} \otimes \dots \otimes A^{(d)}_{k_d} C(k_1, \dots, k_d).$$



Operators in Tensor Train format

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

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



Outline

Parameter-dependent equations

Let us consider a parameter-dependent operator equation

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

where $\xi = (\xi_1, \dots, \xi_s)$ 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).

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 2 (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 $\mathbf{A}$ 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 $\mathbf{b}$ 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

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

and a parameter-dependent equation can also be interpreted as an operator equation on the tensor space $\mathbb{R}^N \otimes \mathbb{R}^{K_1} \otimes \ldots \otimes \mathbb{R}^{K_s}$.

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 3 (Laplace operator)

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

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 4 (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 5 (Laplace operator in tensor train format)

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

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

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 $\{B_{k_{\mathbf{1}},k_{\mathbf{2}}}\}$

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

where

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

Some details about the functional tensor framework

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

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