Christian B. Mendl, Richard M. Milbradt

Tutorial 6 (Solving partial differential equations using a Tucker Ansatz)
Tensor network methods are a powerful tool for solving PDEs which almost "factorize" in coordinate directions. As specific example, we consider Poisson's equation

$$\begin{aligned} -\Delta\phi &= \mathsf{b} \quad \text{on } \Omega = [0,1]^3, \\ \phi|_{\partial\Omega} &= 0 \end{aligned}$$

Here $\Delta = \partial_x^2 + \partial_y^2 + \partial_z^2$ is the Laplace operator, $\mathsf{b} : \Omega \to \mathbb{R}$ is a given function, we search for a solution $\phi : \Omega \to \mathbb{R}$, and $\phi|_{\partial\Omega} = 0$ means that ϕ should be zero on the boundary of Ω (i.e., the faces of the unit cube). We will discretize this PDE and approximate ϕ by a Tucker format representation.[1]

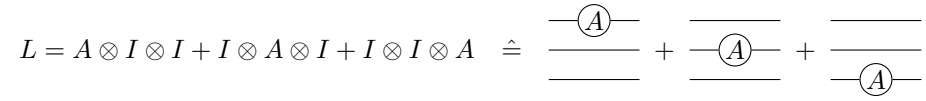
(a) Let $\mathsf{f} : [0,1] \to \mathbb{R}$ be a smooth function. Provide a short derivation of the finite difference approximation

$$-\mathsf{f}''(x) = \frac{-\mathsf{f}(x+h) + 2\mathsf{f}(x) - \mathsf{f}(x-h)}{h^2} + \mathcal{O}(h^2).$$

(b) We now discretize the interval $[0,1]$ using the grid points $x_i = ih$ for $i = 0, 1, \ldots, n$ with $h = \frac{1}{n}$ and some large integer n. Together with zero boundary conditions $\mathsf{f}(0) = \mathsf{f}(1) = 0$, f is discretized as vector $f = (\mathsf{f}(x_1), \ldots, \mathsf{f}(x_{n-1})) \in \mathbb{R}^{n-1}$, omitting the points x_0 and x_n. Assemble a matrix $A \in \mathbb{R}^{(n-1)\times(n-1)}$ such that Af is the discretized negative second derivative of f.

(c) Using such a discretization A of $-\partial_x^2$ in one dimension, show that a discretization of $-\Delta$ on $\Omega = [0,1]^3$ with zero boundary conditions is given by

$L = A \otimes I \otimes I + I \otimes A \otimes I + I \otimes I \otimes A \;\;\hat{=}\;$



where the identity matrix I has the same size as A. Thus $L \in \mathbb{R}^{(n-1)^3 \times (n-1)^3}$.

Now the problem of solving Poisson's equation amounts to solving the linear system

$$L\phi = b. \tag{1}$$

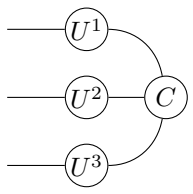
To allow for optimizing a tensor Ansatz for ϕ, we want to convert this to a variational minimization problem. This is achieved by the least squares formulation (using that L is positive semidefinite):

$$\min_{\phi} \frac{1}{2}\langle \phi, L\phi \rangle - \langle \phi, b \rangle \tag{2}$$

(d) Show that (1) and (2) are equivalent when admitting a general vector ϕ, by computing the gradient of (2) with respect to the entries of ϕ. Why must L be positive semidefinite?

To approximate the solution, we use the Tucker format as Ansatz for ϕ, with isometries U^1, U^2, U^3 and a core tensor C:

$\phi =$



with $\quad U^\ell \in \mathbb{R}^{(n-1)\times k_\ell}$ for $\ell = 1, \ldots, 3, \quad C \in \mathbb{R}^{k_1 \times k_2 \times k_3}$

and $k_1, k_2, k_3 \ll n$.

(e) Express $\langle \phi, L\phi \rangle$ and $\langle \phi, b \rangle$ in terms of diagrams, assuming that b is likewise stored in Tucker format.

(f) Specify a corresponding "alternating least squares" (ALS) algorithm to obtain an approximate solution.

[1] In this tutorial all functions and tensors are real-valued.

Solution

(a) We can apply a finite difference scheme for the first derivative twice to obtain the discretized second derivative of f:

$$-f''(x) \approx -\frac{f'(x) - f'(x-h)}{h} \approx -\frac{\frac{f(x+h)-f(x)}{h} - \frac{f(x)-f(x-h)}{h}}{h} = \frac{-f(x+h) + 2f(x) - f(x-h)}{h^2}$$

Alternative derivation based on series $f(x+h) = f(x) + hf'(x) + \frac{1}{2}h^2 f''(x) + \mathcal{O}(h^3)$ possible.)

(b) Matrix-vector form of discretized $f''(x)$, with $f_i = f(x_i)$: Af with

$$A = \frac{1}{h^2}\begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & & \ddots & \\ & & & -1 & 2 \end{pmatrix} \in \mathbb{R}^{(n-1)\times(n-1)}$$

Remark: A is symmetric (clear) and positive semidefinite (proof not discussed here)

(c) Idea: discretized function on unit cube $\Omega = [0,1]^3$ has three indices: ϕ_{i_1,i_2,i_3}, one index for each coordinate. Thus ϕ can be regarded as tensor of degree 3. For the linear system in Eq. (1), we interpret ϕ as long vector with $(n-1)^3$ entries.

Regarding the Laplace operator $\Delta = \partial_x^2 + \partial_y^2 + \partial_z^2$, the derivatives act along each Cartesian coordinate direction separately, and likewise we have to apply A to each dimension of ϕ separately. Written as large matrix, this leads to the representation

$$L = A \otimes I \otimes I + I \otimes A \otimes I + I \otimes I \otimes A,$$

with I the $(n-1) \times (n-1)$ identity matrix. Note that this also requires that we can use the same grid along each Cartesian coordinate, which is the case due to the cubit domain Ω.

(d) To find a possible solution of the minimization problem in (2) we compute the derivative and equate it to zero. The derivative with respect to a single element is:

$$\frac{\mathrm{d}}{\mathrm{d}\phi_k}\left(\frac{1}{2}\sum_{i,j}\phi_i L_{ij}\phi_j - \sum_i \phi_i b_i\right) = \frac{1}{2}\sum_{j\neq k} L_{kj}\phi_j + \frac{1}{2}\sum_{i\neq k}\phi_i L_{ik} + L_{kk}\phi_k - b_k$$

Note that L is symmetric, so

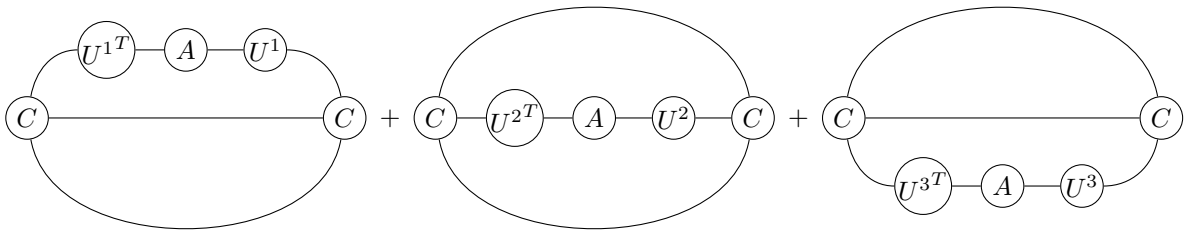
$$\frac{\mathrm{d}}{\mathrm{d}\phi_k}\left(\frac{1}{2}\sum_{i,j}\phi_i L_{ij}\phi_j - \sum_i \phi_i b_i\right) = \sum_i L_{ki}\phi_i - b_k,$$

and therefore

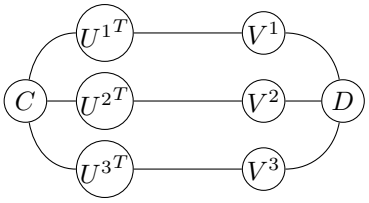
$$\frac{\mathrm{d}}{\mathrm{d}\phi}\left(\frac{1}{2}\langle\phi, L\phi\rangle - \langle\phi, b\rangle\right) = L\phi - b = 0.$$

This gives $L\phi = b$ as in (1). In order for this to be a minimum, L must be positive semidefinite.

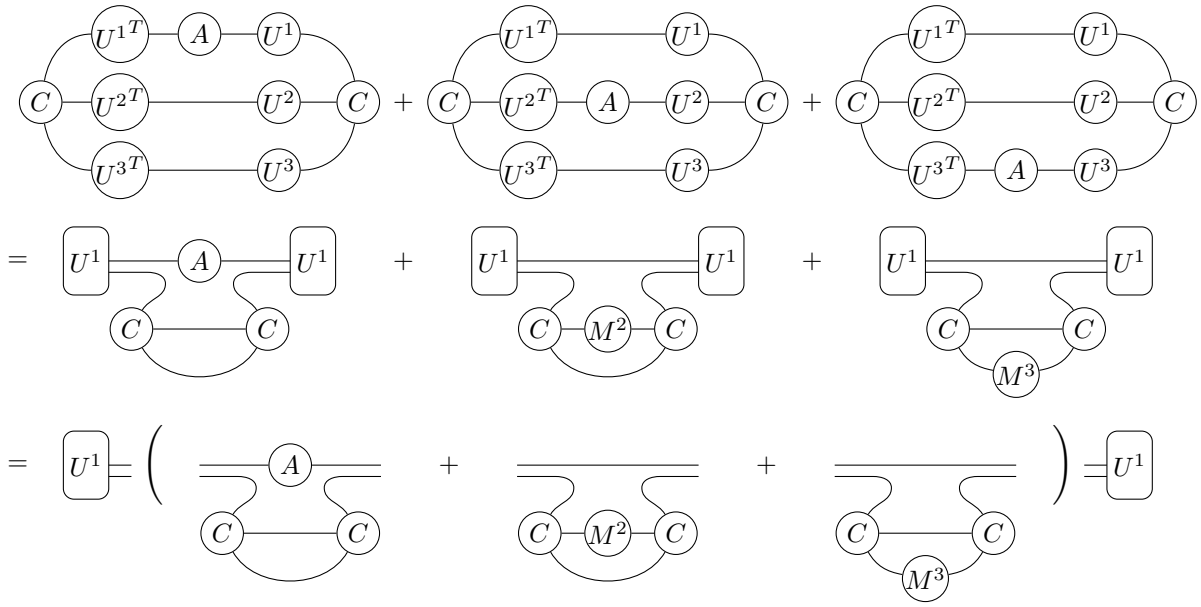
(e) $\langle\phi, L\phi\rangle$:



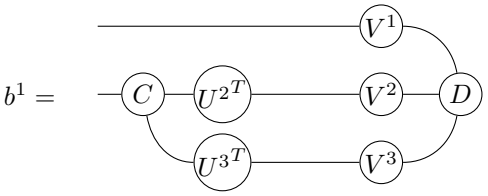
$\langle \phi, b \rangle$:



(f) From the above diagrams one can see that it's possible to split the algorithm into smaller least squares problems. For example, $\min_{U^1} \frac{1}{2}\langle U^1, K^1 U^1 \rangle - \langle U^1, b^1 \rangle$, where K^1 is the operator inside the round brackets here:



and



The same can be done for U^2, U^3 and C, so we can iterate over these optimizations until we reach convergence for ϕ.