Tutorial Machine Learning in Solid Mechanics (Winter term 2022–2023)

TECHNISCHE UNIVERSITÄT DARMSTADT

Cyber-Physical Simulation
TU Darmstadt

Task 3: Hyperelasticity II

Dominik K. Klein, M.Sc. , Yusuf Elbadry, M.Sc. , Prof. Dr. Oliver Weeger
November 29, 2022

In this task you will learn about (i) the trade-off between structure and flexibility and (ii) the *soft* fulfillment of mathematical conditions with NNs.

1 Homogenized cubic lattice metamaterials

1.1 Data preparation

In the first step, data for the calibration and testing of the NN models is to be prepared. Here, synthetic data from homogenized cubic lattice metamaterials will be considered, generated by Fernández et al. [2]. For a short introduction to the dataset, see Klein et al. [3, Section 4.1]. Note that therein the first Piola-Kirchhoff stress is denoted as S, while the deformation gradient is denoted as F. In this tutorial, the first Piola-Kirchhoff stress is denoted as $\boldsymbol{P}$, while the deformation gradient is denoted as $\boldsymbol{F}$. Visit the GitHub repository CPShub/sim-data and download the data for the "BCC" cell. Prepare and visualize the data as in "Task 2: Hyperelasticity I, Section 1" of this tutorial.

For a stable calibration of NNs it is often convenient to use normalization, e.g., to scale the inputs and outputs of the NN in such a way that they are distributed around 1. For the physics-augmented models considered in this tutorial, it is not possible to normalize the NNs input, as this would violate the underlying model physics. However, the output (namely the strain energy and the stress tensor) can be scaled according to

$$W^* = a\,W(\boldsymbol{F})\,, \qquad \boldsymbol{P}^* = a\,\boldsymbol{P}\,, \qquad a \geq 0\,, \tag{1}$$

where W denotes the original strain energy and $\boldsymbol{P}$ denotes the original stress tensor. Scale the data with a suitable factor a so that the components of $\boldsymbol{P}^*$ are in the order of 1.

- Why is it not possible to use batch normalization for this model?
- What are the benefits of scaling the output as in eq. (1)?

1.2 Invariant-based model

In "Task 2: Hyperelasticity I, Section 3" the physics-augmented NN model W^{I} based on invariants was implemented for transversely isotropic materials. Adapt the model for cubic anisotropy by using the vector of invariants

$$\boldsymbol{\mathcal{I}} = (I_1,\, I_2,\, J,\, -J,\, I_7,\, I_{11}) \in \mathbb{R}^6 \tag{2}$$

as input for the model, with the invariants

$$I_1 = \operatorname{tr}\boldsymbol{C}\,, \qquad I_2 = \operatorname{tr}(\operatorname{Cof}\boldsymbol{C}) \qquad J = \det\boldsymbol{F}\,, \qquad I_7 = \boldsymbol{C} : \mathbb{G}_{\text{cub}} : \boldsymbol{C}\,, \qquad I_{11} = \operatorname{Cof}\boldsymbol{C} : \mathbb{G}_{\text{cub}} : \operatorname{Cof}\boldsymbol{C} \tag{3}$$

and the fourth order cubic structural tensor

$$\mathbb{G}_{\text{cub}} = \sum_{i=1}^{3} \boldsymbol{e}_i \otimes \boldsymbol{e}_i \otimes \boldsymbol{e}_i \otimes \boldsymbol{e}_i \,, \tag{4}$$

with $\boldsymbol{e}_i$ being the basis vectors of a Cartesian coordinate system. Use the uniaxial, biaxial, shear and volumetric case for the calibration dataset and the remaining load paths for the test dataset.

- Is the model able to interpolate the dataset? What could be the challenge with this dataset?

2 Deformation-gradient based neural network model

Consider the definition of polyconvexity: The potential $W(\boldsymbol{F})$ is polyconvex if and only if there exists a function $\mathcal{P}\colon \mathbb{R}^{3\times3}\times\mathbb{R}^{3\times3}\times\mathbb{R}\to\mathbb{R}$ such that

$$W(\boldsymbol{F}) = \mathcal{P}(\boldsymbol{F},\, \operatorname{Cof}\boldsymbol{F},\, \det\boldsymbol{F}) \,, \tag{5}$$

so that $\mathcal{P}$ is convex in its arguments [1]. If we use the components of $(\boldsymbol{F},\, \operatorname{Cof}\boldsymbol{F},\, \det\boldsymbol{F}) \in \mathbb{R}^{19}$ as inputs for an ICNN, the scalar-valued output of the ICNN will be polyconvex by construction. However, as we don't use invariants anymore, neither the objectivity condition nor the material symmetry condition are fulfilled by construction.

In hyperelasticity, the objectivity condition is given by

$$W(\boldsymbol{Q}\,\boldsymbol{F}) = W(\boldsymbol{F}) \quad \forall \boldsymbol{F}\in \mathrm{GL}^+(3),\; \boldsymbol{Q}\in \mathrm{SO}(3)\,, \tag{6}$$

which implies the stress invariance condition

$$\boldsymbol{P}(\boldsymbol{Q}\,\boldsymbol{F}) = \boldsymbol{Q}\,\boldsymbol{P}(\boldsymbol{F}) \quad \forall \boldsymbol{F}\in \mathrm{GL}^+(3),\; \boldsymbol{Q}\in \mathrm{SO}(3)\,. \tag{7}$$

Here,

$$\mathrm{GL}^+(3)\colon = \left\{\boldsymbol{X}\in\mathbb{R}^{3\times3} \,|\, \det\boldsymbol{X} > 0\right\} \tag{8}$$

denotes the set of invertible second order tensors with positive determinant, while

$$\mathrm{SO}(3) := \left\{\boldsymbol{X}\in\mathbb{R}^{3\times3} | \boldsymbol{X}\boldsymbol{X}^T = \boldsymbol{I},\, \det\boldsymbol{X} = 1\right\} \tag{9}$$

denotes the special orthogonal group in $\mathbb{R}^3$. Simply said, the $\mathrm{SO}(3)$ contains all rotation matrices in the three-dimensional space without inversion.

The material symmetry condition is given by

$$W(\boldsymbol{F}\,\boldsymbol{Q}) = W(\boldsymbol{F}) \quad \forall \boldsymbol{F}\in \mathrm{GL}^+(3),\, \boldsymbol{Q}\in \mathcal{G}\subseteq \mathrm{SO}(3)\,. \tag{10}$$

Here, $\mathcal{G}$ denotes the symmetry group of the material under consideration. This implies the stress invariance condition

$$\boldsymbol{P}(\boldsymbol{F}\,\boldsymbol{Q}) = \boldsymbol{P}(\boldsymbol{F})\,\boldsymbol{Q} \quad \forall \boldsymbol{F}\in \mathrm{GL}^+(3),\, \boldsymbol{Q}\in \mathcal{G}\subseteq \mathrm{SO}(3)\,. \tag{11}$$

For the cubic group $\mathcal{G}_7$ (see [1, Sec.3.7]), which we consider in this task, the 24 elements of the symmetry group are given by

- (1) the identity matrix $\boldsymbol{I}$,
- (2-4) rotations of $\{\pi/2,\, \pi,\, 3\pi/2\}$ around the x-axis,
- (5-7) rotations of $\{\pi/2,\, \pi,\, 3\pi/2\}$ around the y-axis,
- (8-10) rotations of $\{\pi/2,\, \pi,\, 3\pi/2\}$ around the z-axis,
- (11-16) rotations of π around the axes $(1,1,0)$, $(-1,1,0)$, $(1,0,1)$ $(-1,0,1)$, $(0,1,1)$, $(0,-1,1)$,
- (17-20) rotations of $2\pi/3$ around the axes $(1,1,1)$, $(-1,1,1)$, $(1,-1,1)$, $(-1,-1,1)$
- (21-24) and rotations of $4\pi/3$ around the axes $(1,1,1)$, $(-1,1,1)$, $(1,-1,1)$, $(-1,-1,1)$.

Adapt the model of "Task 2: Hyperelasticity I, Sec. 3" so that it takes the components of $(\boldsymbol{F},\, \operatorname{Cof}\boldsymbol{F},\, \det\boldsymbol{F}) \in \mathbb{R}^{19}$ as input. We denote this model as W^{F}. Calibrate the model using the calibration dataset defined in Sec. 1.2. Implement an evaluation of W^{F} which evaluates the model for multiple observers as well as for all material symmetry operations. For the multiple observers, choose a finite amount of randomly distributed rotation matrices. Check this evaluation using the invariant-based model W^{I}. Then, for the calibrated model W^{F}, evaluate the invariance conditions on both the calibration and test dataset and find a suitable visualization for the results of the different observers and symmetry transformations.

- How does the polyconvexity condition restrict the first hidden layer of the NN when the components of $(\boldsymbol{F},\, \operatorname{Cof}\boldsymbol{F},\, \det\boldsymbol{F}) \in \mathbb{R}^{19}$ are used as inputs? What is the difference to the invariant-based NN model from "Task 2: Hyperelasticity I" of this tutorial?
- Which physical / mathematical conditions are fulfilled by this model?
- We have to check two invariance conditions. Which possibilities exist to examine them both? (hint: concurrent or succesively)

3 Data augmentation

The model W^{F} does not fulfill the objectivity condition and material symmetry condition by construction. Now, data augmentation is used to "teach" the model how to fulfill both conditions in an approximate fashion. Note that, for finite symmetry groups such as the cubic one, it is also possible to fulfill the material symmetry group in an exact way, c.f. [2]. For the data augmentation, based on the initial dataset $D = \{\boldsymbol{F},\, W,\, \boldsymbol{P}\}$, the dataset is extended by a finite amount of artificial observers and by the elements of the symmetry group according to

$$\widetilde{D} = \bigcup_{\substack{\boldsymbol{Q}_{\text{obj}} \in \mathcal{G}_{\text{obj}}, \\ \boldsymbol{Q}_{\text{mat}} \in \mathcal{G}_{\text{mat}}}} \left\{ \boldsymbol{Q}_{\text{obj}}\, \boldsymbol{F}\, \boldsymbol{Q}_{\text{mat}},\, W,\, \boldsymbol{Q}_{\text{obj}}\, \boldsymbol{P}\, \boldsymbol{Q}_{\text{mat}} \right\} \, . \tag{12}$$

Here, $\mathcal{G}_{\text{mat}}$ denotes the symmetry group under consideration and $\mathcal{G}_{\text{obj}} \subset \mathrm{SO}(3)$ denotes a finite amount of randomly distributed rotation matrices. While the number of elements in the cubic symmetry group is fixed, the number of rotation matrices in $\mathcal{G}_{\text{obj}}$ for the objectivity condition has to be chosen.

Augment the initial calibration dataset and calibrate the model W^{F}. Increase the number of rotation matrices included in $\mathcal{G}_{\text{obj}}$ in $[8,\, 16,\, 32,\, \dots]$ until the objectivity condition is approximated "well enough". For this, use different approaches (concurrent / succesively) to augment the dataset. Hint: it is useful to pre-calibrate the model W^{F} with only one observer before using the larger, augmented dataset.

- How must the test dataset be chosen to check the fulfillment of the objectivity condition and the material symmetry condition?
- Out of all datasets and models you used in this tutorial – which one would you prefer in each case?

Optional Calibrate the model W^{F} to the transversely isotropic data of "Task 2: Hyperelasticity I, Sec. 1". The transversely isotropic symmetry group is given by

$$\mathcal{G}_{\text{ti}} := \{\boldsymbol{Q}_{\boldsymbol{x}}^{\alpha} \,|\, \alpha \in [0,\, 2\pi]\} \ , \tag{13}$$

where $\boldsymbol{Q}_{\boldsymbol{x}}^{\alpha}$ denotes a rotation around the preferred direction $\boldsymbol{x}$ by the angle α. In this example, the preferred direction is the x-axis. Use different amounts of equidistant elements of this group for the data augmentation.

References

[1] V. Ebbing. "Design of Polyconvex Energy Functions for All Anisotropy Classes". PhD thesis. Universität Duisburg-Essen, 2010.

[2] M. Fernández, M. Jamshidian, T. Böhlke, K. Kersting and O. Weeger. "Anisotropic hyperelastic constitutive models for finite deformations combining material theory and data-driven approaches with application to cubic lattice metamaterials". In: *Computational Mechanics* 67.2 (2021), pp. 653–677. DOI: 10.1007/s00466-020-01954-7.

[3] D. K. Klein, M. Fernández, R. J. Martin, P. Neff and O. Weeger. "Polyconvex anisotropic hyperelasticity with neural networks". In: *Journal of the Mechanics and Physics of Solids* 159 (2022), p. 104703. DOI: 10.1016/j.jmps.2021.104703. arXiv: 2106.14623.