

Оглавление

1	CLaNN	2
1.1	Введение	2
1.2	Кинематика	2
1.3	Напряжение и термодинамическая корректность	3
1.4	Архитектура CLaNN и её производные	4
1.5	Виртуальный эксперимент	8
1.5.1	Интерполяция и экстраполяция кривых нагружения	11
1.5.2	Виртуальные эксперименты	11
1.6	Заключение	11
A	Эквивалентность QR-факторизации \mathbf{F} и разложения Холецкого $\mathbf{C}=\mathbf{F}^\top\mathbf{F}$ для вычисления логарифмических координат $\boldsymbol{\xi}$	14
A.1	Постановка и обозначения	14
A.2	Теорема (эквивалентность \mathbf{U} и \mathbf{R})	14
A.3	Координаты $\boldsymbol{\xi}$ через \mathbf{U}	15

Глава 1

CLANN

1.1 Введение

В данной работе представлена термодинамически корректная гиперупругая модель CLaNN (Convex Laplace Neural Network), основанная на выпуклой нейросети и логарифмической параметризации деформации - тензоре Лапласа. Модель обеспечивает физически корректное интерполирование и экстраполирование поля напряжений и объективное описание механического поведения материалов при больших деформациях, что было проверено при помощи обучения модели на синтетических и натуральных данных, и валидации получившейся модели на численном эксперименте.

1.2 Кинематика

Основные соотношения

Мы рассматриваем равновесие тонкой несжимаемой гиперупругой мембраны под определенными нагрузками. Деформация мембраны характеризуется деформацией её поверхности. Обозначим через \mathbf{X} и \mathbf{x} положения точек, в соответствующих базисах \mathbf{E}_α и \mathbf{e}_α , в исходной (недеформированной) $\Omega_0 \subset \mathbb{R}^2$ и текущей (деформированной) $\Omega_t \subset \mathbb{R}^2$ конфигурациях поверхности мембраны соответственно. Деформация определяется отображением $\mathbf{x} = \mathbf{x}(\mathbf{X})$, поверхностный градиент деформации $\mathbf{F} = \mathbf{e}_\alpha \otimes \mathbf{E}^\alpha$, а правый тензор Коши—Грина $\mathbf{C} = C_{\alpha\beta} \mathbf{e}_\alpha \otimes \mathbf{e}_\beta = \mathbf{F}^\top \mathbf{F}$. Для определения меры деформации мы используем меру Лапласа $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3)^T$ [9], которая может быть вычислен двумя эквивалентными способами: либо через QR-разложение градиента деформации $\mathbf{F} = \mathbf{Q}\mathbf{R}$ с $\mathbf{U} = \mathbf{R}$, либо через разложение Холецкого правого тензора Коши-Грина $\mathbf{C} = \mathbf{U}^\top \mathbf{U}$ (Приложение А). В этом случае гиперупругий потенциал является функцией от деформации Лапласа $\psi = \psi(\boldsymbol{\xi})$.

Мера деформации Лапласа В двумерном случае вводятся характеристики

$$\xi_1 = \ln(u_{11}), \quad \xi_2 = \ln(u_{22}), \quad \xi_3 = \frac{u_{12}}{u_{11}}, \quad \mathbf{U} = u_{\alpha\beta} \mathbf{e}_\alpha \otimes \mathbf{e}_\beta. \quad (1.1)$$

1.3 Напряжение и термодинамическая корректность

Второй тензор напряжений Пиолы-Кирхгофа вычисляется по цепному правилу дифференцированием энергии ψ по правому тензору деформации Коши-Грина \mathbf{C} :

$$\mathbf{S} = 2 \frac{\partial \psi}{\partial \mathbf{C}} = 2 \frac{\partial \psi}{\partial \boldsymbol{\xi}} \cdot \frac{\partial \boldsymbol{\xi}}{\partial \mathbf{C}} = 2 \mathbf{r}(\boldsymbol{\xi}) \cdot \frac{\partial \boldsymbol{\xi}}{\partial \mathbf{C}}, \quad \mathbf{r} := \frac{\partial \psi}{\partial \boldsymbol{\xi}}. \quad (1.2)$$

Такое построение имеет ключевые следствия:

- **Объективность:** $\psi(\mathbf{C}) = \psi(\mathbf{Q}^\top \mathbf{C} \mathbf{Q})$ для любой ортогональной \mathbf{Q} , а значит и \mathbf{S} инвариантен к поворотам.
- **Симметрия напряжений:** $\mathbf{S} = \mathbf{S}^\top$ вследствие симметрии \mathbf{C} и корректного применения цепного правила.
- **Термодинамическая корректность:** равенство (1.2) является следствием неравенства Клаузиуса-Дюгема $\mathcal{D} = \mathbf{S} : \dot{\mathbf{C}} - \dot{\psi}(\mathbf{C}) \geq 0$, выражающее второе начало термодинамики для механических процессов [10, 11].

Связь тензора Лапласа и второго тензора напряжений Пиолы-Кирхгофа

Применяя цепное правило дифференцирования к выражению (1.2) и используя меру деформации Лапласа, получаем аналитические выражения для компонент второго тензора напряжений Пиолы-Кирхгофа в двумерном случае:

$$\begin{aligned} S_{11} &= e^{-2\xi_1} (r_1 - 2\xi_3 r_3) + e^{-2\xi_2} r_2 \xi_3^2, \\ S_{22} &= e^{-2\xi_2} r_2, \\ S_{12} &= -e^{-2\xi_2} r_2 \xi_3 + e^{-2\xi_1} r_3, \end{aligned} \quad (1.3)$$

r_1, r_2, r_3 — компоненты функции отклика $\mathbf{r} = \frac{\partial \psi}{\partial \boldsymbol{\xi}}$.

Эти соотношения демонстрируют связь между логарифмическими мерами деформации и компонентами напряжений, характерную для гиперупругих материалов. Экспоненциальные множители $e^{-2\xi_i}$ отражают логарифмическую природу выбранной параметризации, а члены, содержащие ξ_3 , описывают сдвиговые эффекты.

Фундаментальные ограничения

В соответствии с принципами термодинамики и механики сплошных сред, гиперупругая модель должна удовлетворять ряду фундаментальных ограничений, обеспечивающих физическую корректность и материальную устойчивость.

Положительность и рост энергии деформации:

$$\psi(\boldsymbol{\xi}) \geq 0 \quad \forall \boldsymbol{\xi} \in \mathbb{R}^3, \quad \psi(\boldsymbol{\xi}) \rightarrow 0 \text{ при } \|\boldsymbol{\xi}\| \rightarrow 0 \quad \psi(\boldsymbol{\xi}) \rightarrow \infty \text{ при } \|\boldsymbol{\xi}\| \rightarrow \infty. \quad (1.4)$$

Эти свойства принято записывать через градиент деформации \mathbf{F} и правый

тензор деформации Коши-Грина \mathbf{C} [2, 4, 6], но они эквивалентны и для меры деформации Лапласа $\boldsymbol{\xi}$.

Напряжения в естественной конфигурации

Из условия $\psi(\boldsymbol{\xi}) \rightarrow 0$ при $\|\boldsymbol{\xi}\| \rightarrow 0$ и непрерывности функции энергии следует, что в естественном состоянии напряжения также обращаются в нуль:

$$\mathbf{S} \rightarrow \mathbf{0} \text{ при } \|\boldsymbol{\xi}\| \rightarrow 0, \quad (1.5)$$

Это условие является прямым следствием определяющего соотношения $\mathbf{S} = 2\partial\psi/\partial\mathbf{C}$ и принципа объективности, согласно которому в отсутствие деформации не может быть внутренних напряжений.

1.4 Архитектура CLaNN и её производные

В рамках предложенного под хода CLaNN (Convex Laplace Neural Network) энергия деформации $\psi(\boldsymbol{\xi})$ с мерой деформации Лапласа аппроксимируется посредством выпуклой по входу нейронной сетью (Input Convex Neural Network, ICNN) [1] и вычисления 2 тензора напряжения Пиолы-Кирхгофа \mathbf{S} .

Обобщенная архитектура ICNN

ICNN представляет собой класс нейронных сетей, гарантирующих выпуклость выходной функции относительно входных переменных. Функция $\psi : \mathbb{R}^3 \rightarrow \mathbb{R}$ называется выпуклой, если $\forall \boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in \mathbb{R}^3$ и $\lambda \in [0, 1]$ выполняется неравенство Йенсена:

$$\psi(\lambda \boldsymbol{\xi}_1 + (1 - \lambda) \boldsymbol{\xi}_2) \leq \lambda \psi(\boldsymbol{\xi}_1) + (1 - \lambda) \psi(\boldsymbol{\xi}_2). \quad (1.6)$$

Общая архитектура ICNN с L скрытыми слоями определяется следующими соотношениями:

$$z_1 = \sigma(\mathbf{W}_1 \boldsymbol{\xi} + \mathbf{b}_1), \quad (1.7)$$

$$z_{i+1} = \sigma(\mathbf{W}_{i+1} z_i + \mathbf{U}_{i+1} \boldsymbol{\xi} + \mathbf{b}_{i+1}), \quad i = 1, \dots, L - 1, \quad (1.8)$$

$$\psi = \mathbf{W}_{L+1} z_L + \mathbf{U}_{L+1} \boldsymbol{\xi} + b_{L+1}, \quad (1.9)$$

где σ — выпуклая и неубывающая функция активации (например, ReLU, softplus), $\mathbf{W}_i \geq 0$ — неотрицательные весовые матрицы скрытых слоев, \mathbf{U}_i — произвольные весовые матрицы для прямых связей от входа, \mathbf{b}_i — векторы смещений. Ключевые принципы построения ICNN включают: (1) использование неотрицательных весов в скрытых слоях, (2) применение выпуклых функций активации, (3) линейные связи от входного слоя ко всем последующим слоям.

В данной работе используется упрощенная однослойная архитектура ICNN с одним скрытым слоем, определяемая следующими соотношениями:

$$s = \mathbf{W}_1 \boldsymbol{\xi} + \mathbf{b}_1, \quad z = \frac{\text{softplus}(\beta s)}{\beta}, \quad z_0 = \frac{\text{softplus}(\beta \mathbf{b}_1)}{\beta}, \quad \psi = \mathbf{W}_2^\top (z - z_0), \quad \mathbf{W}_2 \geq 0. \quad (1.10)$$

Калибровка естественного состояния и строгая неотрицательность. Определим $\mathbf{r}_0 := \partial\psi/\partial\boldsymbol{\xi}|_{\boldsymbol{\xi}=\mathbf{0}}$. В реализации используем откорректированный потенциал

$$\psi_{\text{phys}}(\boldsymbol{\xi}) = \mathbf{W}_2^\top (z - z_0) - \mathbf{r}_0^\top \boldsymbol{\xi}, \quad (1.11)$$

что обеспечивает $\psi_{\text{phys}}(\mathbf{0}) = 0$, $\mathbf{S}(\mathbf{I}) = \mathbf{0}$ и строгую неотрицательность $\psi_{\text{phys}}(\boldsymbol{\xi}) > 0$ при $\boldsymbol{\xi} \neq \mathbf{0}$.

Функция softplus определяется как $\text{softplus}(x) = \ln(1 + e^x)$ [3]. Функция активации $\text{softplus}(\beta x)/\beta$ строго выпукла и $\lim_{\beta \rightarrow \infty} \text{softplus}(\beta x)/\beta = \max(0, x)$ (ReLU). Условие $\mathbf{W}_2 \geq 0$ сохраняет выпуклость линейной комбинации.

Центровка энергии в естественном состоянии: Ключевой особенностью предложенной архитектуры является гарантированное выполнение условия $\psi(\mathbf{0}) = 0$ через вычитание $z_0 = \text{softplus}(\beta \mathbf{b}_1)/\beta$. Это обеспечивает физически корректное поведение модели в естественном (недеформированном) состоянии, где энергия деформации должна обращаться в нуль, что является фундаментальным требованием гиперупругой теории.

Доказательство зануления энергии: При $\boldsymbol{\xi} = \mathbf{0}$ имеем $s(\mathbf{0}) = \mathbf{b}_1$, откуда $z(\mathbf{0}) = z_0$ и $\psi(\mathbf{0}) = \mathbf{W}_2^\top (z(\mathbf{0}) - z_0) = 0$, где $\mathbf{W}_2^+ = |\mathbf{W}_2| \geq 0$. Поскольку z_0 не зависит от $\boldsymbol{\xi}$, производные $\partial\psi/\partial\boldsymbol{\xi}$ и $\partial^2\psi/\partial\boldsymbol{\xi}^2$ остаются неизменными, сохраняя корректность напряжений и гессиана энергии.

После прохождения через все ребра вычислительного графа и получения ψ происходит автоматическое дифференцирование $\partial\psi/\partial\boldsymbol{\xi}$, которое реализована во всех современных библиотеках для машинного обучения, после чего вычисляется тензор напряжений \mathbf{S} по формуле (1.3).

Размерности весовых параметров сети: $\mathbf{W}_1 \in \mathbb{R}^{h \times 3}$, $\mathbf{b}_1 \in \mathbb{R}^h$, $\mathbf{W}_2 \in \mathbb{R}_{\geq 0}^h$, где h — размерность скрытого слоя. Отсутствие дополнительного смещения b_2 обеспечивает выполнение условия $\psi(\mathbf{0}) = 0$.

Аналитические выражения для производных энергии

Градиент энергии деформации

Аналитическое дифференцирование функции энергии по переменным $\boldsymbol{\xi}$ даёт выражение для градиента:

$$\nabla_{\boldsymbol{\xi}} \psi = \mathbf{W}_1^T (\mathbf{W}_2 \odot \sigma(\beta(\mathbf{W}_1 \boldsymbol{\xi} + \mathbf{b}_1))), \quad (1.12)$$

где $\sigma(x) = \frac{1}{1+e^{-x}}$ — сигмоида, а операция \odot обозначает поэлементное произведение (Hadamard product). Данное выражение демонстрирует, что градиент энергии является линейной комбинацией строк матрицы \mathbf{W}_1^T с весами, определяемыми произведением выходных весов \mathbf{W}_2 и значений функции активации $\sigma(\beta(\mathbf{W}_1 \boldsymbol{\xi} + \mathbf{b}_1))$.

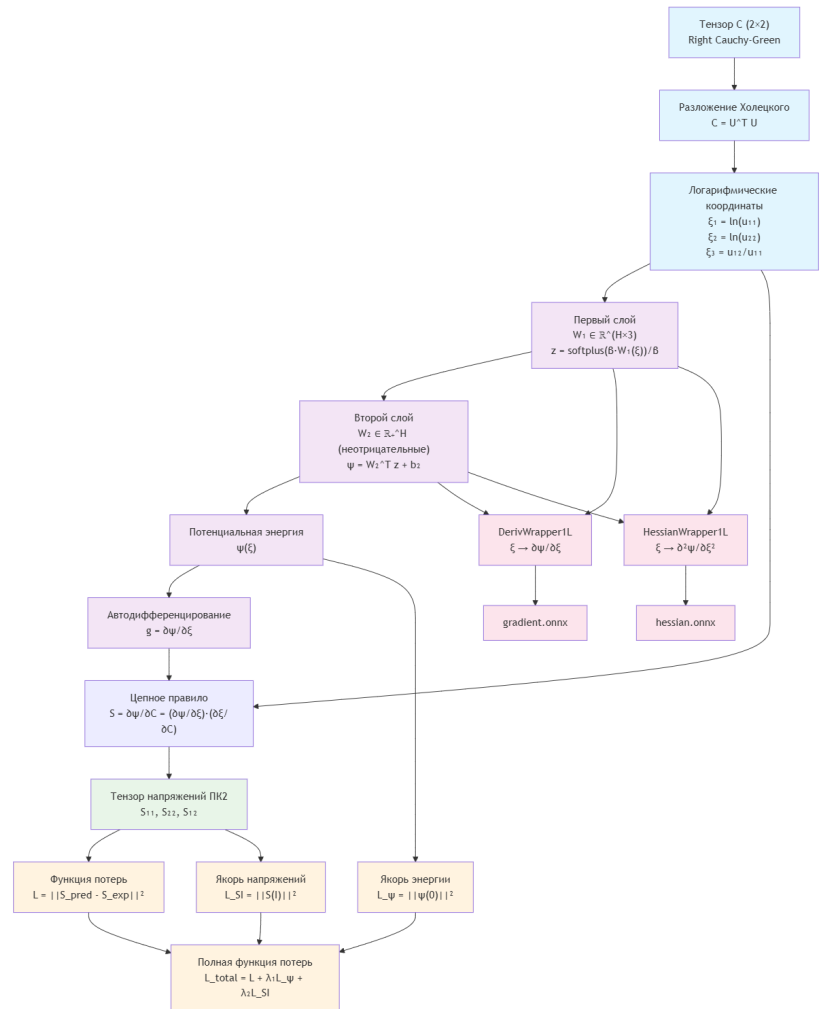


Рис. 1.1. Схема вычислительного процесса CLANN: от входного тензора до функции потерь. Показаны этапы обработки входных данных, вычисления нейросетью, дифференцирования и формирования функции потерь.

Гессиан энергии деформации

Вторые производные энергии по переменным ξ определяют гессиан, который имеет следующую аналитическую форму:

$$H_{ij} = \sum_h \sigma'_h W_{2,h} W_{h,i} W_{h,j}, \quad (1.13)$$

где $\sigma' = \beta \sigma(1 - \sigma)$ - производная сигмоиды, $\sigma = \text{sigmoid}(\beta s)$, а $s = \mathbf{W}_1 \xi + \mathbf{b}_1$.

Материальная устойчивость и положительная определённость

Из строгой выпуклости $\psi(\xi)$ следует положительная определённость гессиана:

$$\mathbf{H} = \frac{\partial^2 \psi}{\partial \xi^2} > 0, \quad (1.14)$$

что обеспечивает положительную определённость касательных модулей упругости $\mathbb{C} = \partial^2 \psi / \partial \mathbf{C}^2$ через цепное правило дифференцирования. Это свойство важно для численной стабильности конечно-элементных расчётов, поскольку на практике обеспечивает сходимость метода Ньютона и отсутствие сингулярностей в матрице жёсткости.

В рамках предложенного подхода обучение модели осуществляется путём минимизации функции потерь, которая количественно характеризует невязку между предсказанными и экспериментальными значениями напряжений:

$$L = \frac{1}{N} \sum_{i=1}^N \|\mathbf{S}_{\text{pred}}^{(i)} - \mathbf{S}_{\text{exp}}^{(i)}\|^2. \quad (1.15)$$

Для минимизации функции потерь (1.15) используется оптимизатор Adam [5], который широко используется в задачах машинного обучения. Процесс оптимизации включает вычисление градиентов по всем параметрам сети и обновление весов с использованием адаптивных моментов первого и второго порядка.

Такое построение архитектуры CLaNN обеспечивает выполнение всех необходимых физических свойств гиперупругой модели: **термодинамическая корректность** достигается через строгое соблюдение соотношения (1.2), что гарантирует консервативность напряжений $\oint \mathbf{S} : d\mathbf{C} = 0$ и согласованность с законами термодинамики; **материальная устойчивость** обеспечивается и существенно улучшается за счёт строгой выпуклости функции энергии $\psi(\xi)$, гарантируемой архитектурой ICNN ($\mathbf{W}_2 \geq 0$, выпуклая неубывающая активация); **объективность** автоматически выполняется благодаря параметризации через тензор Коши-Грина $\mathbf{C} = \mathbf{F}^\top \mathbf{F}$, обеспечивая инвариантность относительно поворотов и симметрию напряжений; **строгая неотрицательность и коэрцитивность энергии** обеспечиваются архитектурной калибровкой $\psi_{\text{phys}}(\xi) = \mathbf{W}_2^\top (z - z_0) - \mathbf{r}_0^\top \xi + \frac{\gamma}{2} \|\xi\|^2$ с $\gamma > 0$, что даёт $\psi_{\text{phys}}(\mathbf{0}) = 0$, $\psi_{\text{phys}}(\xi) > 0$ при $\xi \neq \mathbf{0}$ и $\psi_{\text{phys}}(\xi) \rightarrow \infty$ при $\|\xi\| \rightarrow \infty$ (см. clann_math.markdown, § Потенциал); **численная стабильность** достигается через логарифмическую параметризацию Лапласа (1.1), которая корректно обрабатывает большие деформации

и предотвращает нефизичные значения; наконец, **физические ограничения** (1.4) обеспечиваются архитектурой сети CLaNN: монотонные, выпуклые функции активации, неотрицательные весовые коэффициенты, а также нормировка нулевого уровня энергии: из потенциальной энергии вычитается её значение в естественной конфигурации, что обеспечивает $\psi(0) = 0$ и $\mathbf{S}(\mathbf{I}) = \mathbf{0}$.

1.5 Виртуальный эксперимент

Мы используем синтетические экспериментальные данные для тестирования CLaNN на изотропном надувании однородной и неоднородной по толщине мембраны. А именно, мы генерируем данные с помощью виртуальных экспериментов на плоских растяжениях образца и используем их в качестве входных данных для обучения CLaNN, без какого-либо дополнительного знания об изотропности/анизотропии образца и форме потенциала.

Обучение модели проводилось на численных экспериментальных данных, полученных при двухосном растяжении образца с геометрией мальтийского креста (Рисунок 1.2). С неогуковской гиперупругой моделью для двумерной мембраны [8]

Причем данные для обучения собирались из одного центрального элемента сетки, что соответствует ограничениям эквивалентного натурного эксперимента, в котором невозможно установить без предположения модели материала полное поле напряжения в образце.

Для решения задачи равновесия гиперупругой мембраны используется метод описанный в [7].

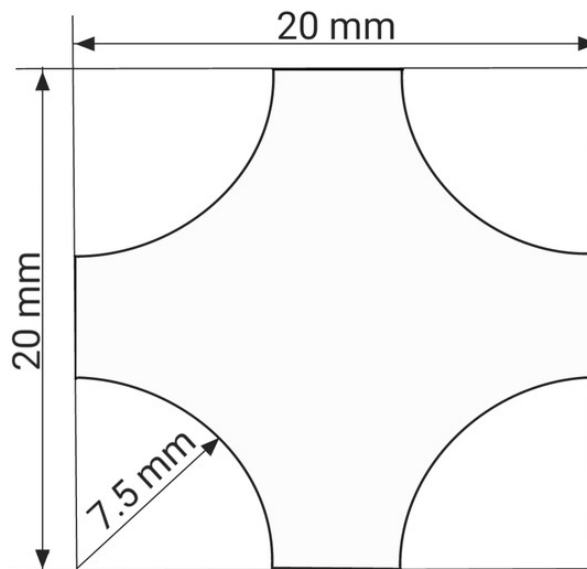


Рис. 1.2. Размеры образца биоматериала в форме мальтийского креста. Радиус вырезов одинаков для всех вырезов

Схематическое представление протоколов показано на рисунке 1.3, где $w_i \in [0, 1]$, $i \in \{1, 4\}$ представляет собой долю от заданного максимального смещения u_{\max}

для i -го плеча: $w_i = 0$ соответствует неподвижному плечу, а $w_i = 1$ — соответствует плечу, чьё положение было сдвинуто и фиксировано на расстояние u_{\max} . Изменяя w_i , можно получить различные типы экспериментов. В наших виртуальных экспериментах мы постепенно прикладываем смещение с определенным шагом Δs до достижения максимального смещения. Смещение $w_i \cdot n \cdot \Delta s$ прикладывается к i -му плечу на n -м шаге, где $n = 1, \dots, N$, $N = u_{\max}/\Delta s$ — количество шагов. Треугольная сетка для образца является квазиравномерной с размером ячейки $h_{\text{fit}} = 0.25$ мм, максимальное смещение $u_{\max} = 2$ мм и $\Delta s = 0.2$ мм. На каждом шаге мы собираем данные (\mathbf{C}, \mathbf{S}) для всех треугольников, принадлежащих центральной области. Поскольку мы используем линейные (P_1) конечные элементы, значения (\mathbf{C}, \mathbf{S}) постоянны на каждом треугольнике.

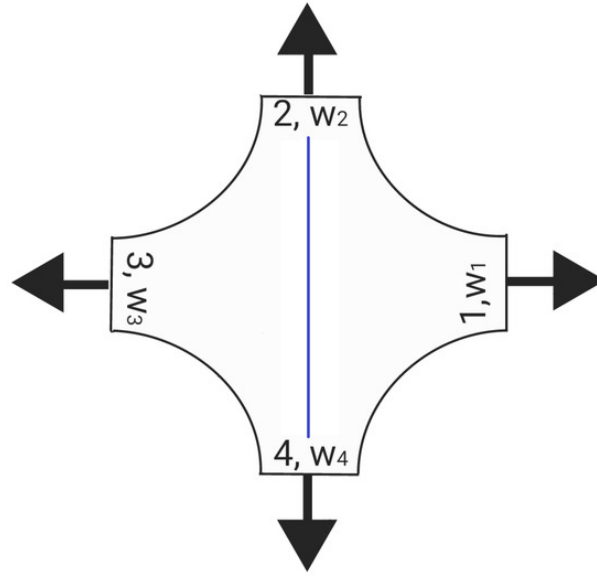


Рис. 1.3. Схематическое представление протоколов. Радиус вырезов одинаков для всех вырезов

Наш предлагаемый тестовый протокол предполагает девять экспериментов:

Таблица 1.1. Протоколы тестовых экспериментов

№	w_1	w_2	w_3	w_4
1	1	1	1	1
2	1	0.75	1	0.75
3	0.75	1	0.75	1
4	1	0.5	1	0.5
5	0.5	1	0.5	1
6	1	1/3	1	1/3
7	1/3	1	1/3	1
8	1	0	1	0
9	0	1	0	1

Обучающий набор содержит 90 точек данных правого тензора деформаций Коши-Грина \mathbf{C} и второго тензора напряжений Пиолы-Кирхгофа \mathbf{S} .

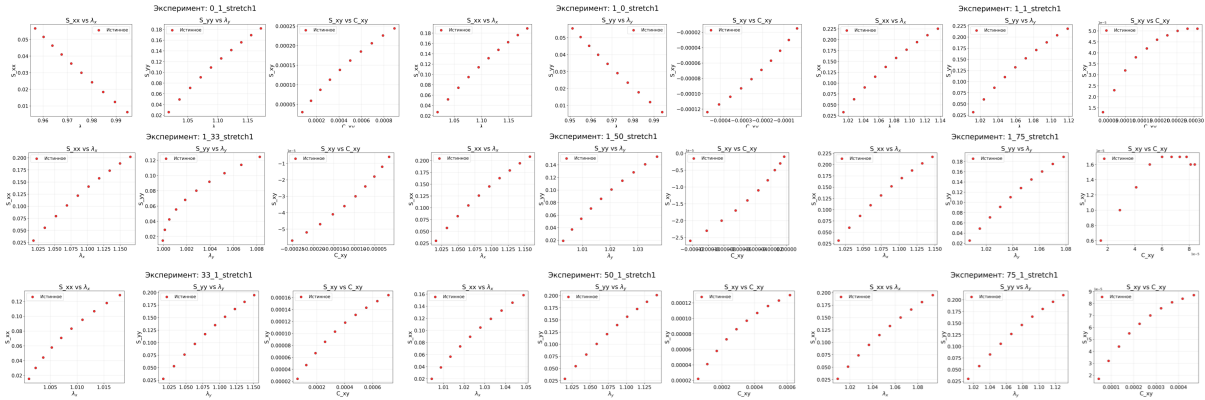


Рис. 1.4. Обучающий набор данных

Гиперпараметры оптимизации:

- Скорость обучения (learning rate): 0.001
- Размер батча (batch size): 128
- Архитектура: 16 нейронов на скрытом слое

Результаты обучения: Процесс оптимизации показал высокую эффективность: ошибка аппроксимации снизилась на 5 порядков за менее чем 5000 эпох (рис. 1.5), что демонстрирует как качество предложенной архитектуры, так и корректность выбора функции потерь. Столь быстрая сходимость обусловлена строгой выпуклостью функции энергии, что обеспечивает единственность минимума и отсутствие локальных минимумов в пространстве параметров.

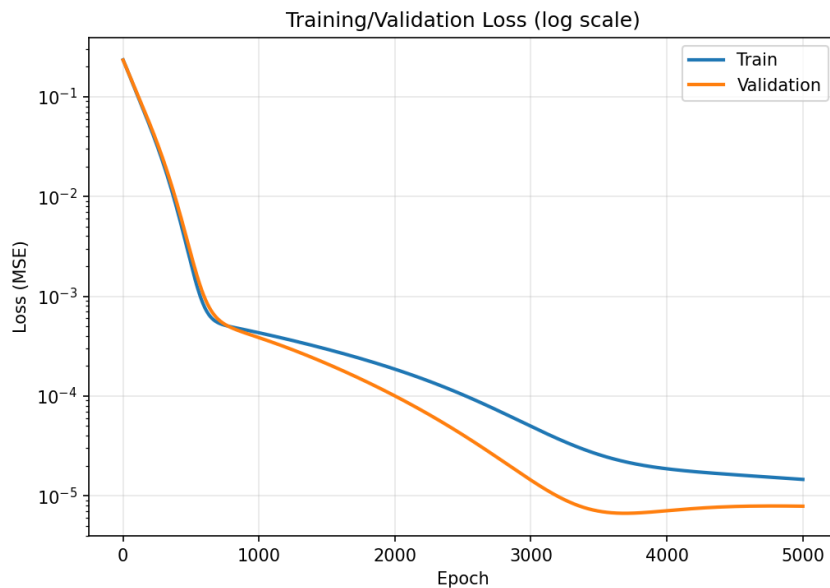


Рис. 1.5. Кривая функции потерь при обучении на 90 точках данных

1.5.1 Интерполяция и экстраполяция кривых нагружения

Сначала мы проверили, как модель CLANN интерполирует и экстраполирует кривые нагружения. Для этого мы использовали 90 точек набора данных и разбивали его на различные соотношения обучающего и тестового наборов. Мы построили кривые нагружения на каждый протокол и сравнили их с экспериментальными данными. Для этого использовали коэффициент детерминации R^2 , определяемый как:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (1.16)$$

где y_i — экспериментальные значения, \hat{y}_i — предсказания модели, \bar{y} — среднее экспериментальных значений, n — количество точек данных. Результаты представлены на рисунке 1.6.

Для каждого протокола мы построили кривую нагружения и сравнили её с экспериментальными данными.

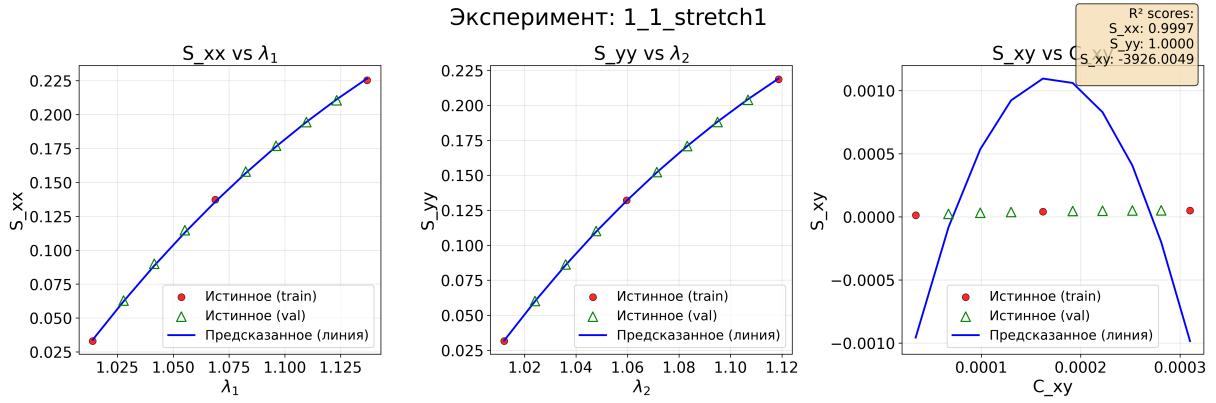


Рис. 1.6. Кривая нагружения для равнодвухосного растяжения

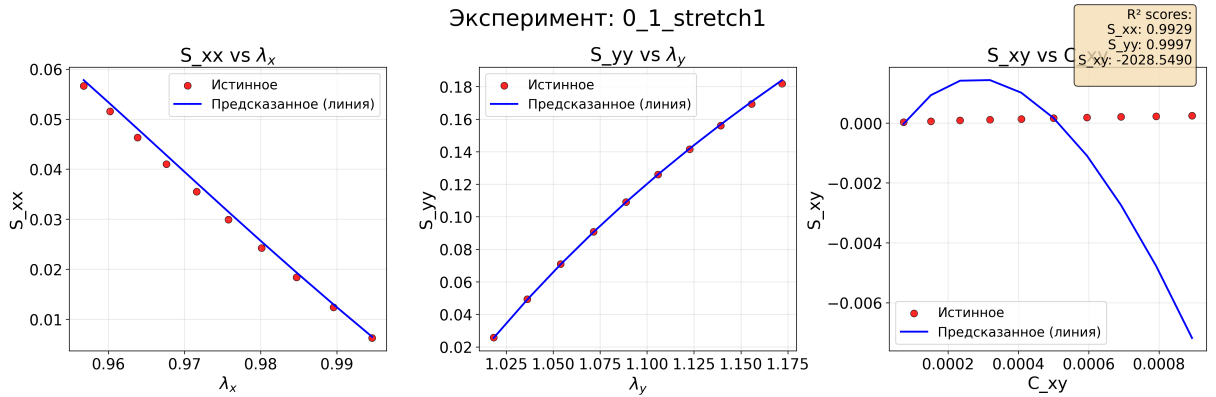


Рис. 1.7. Кривая нагружения для неравнодвухосного растяжения

1.5.2 Виртуальные эксперименты

1.6 Заключение

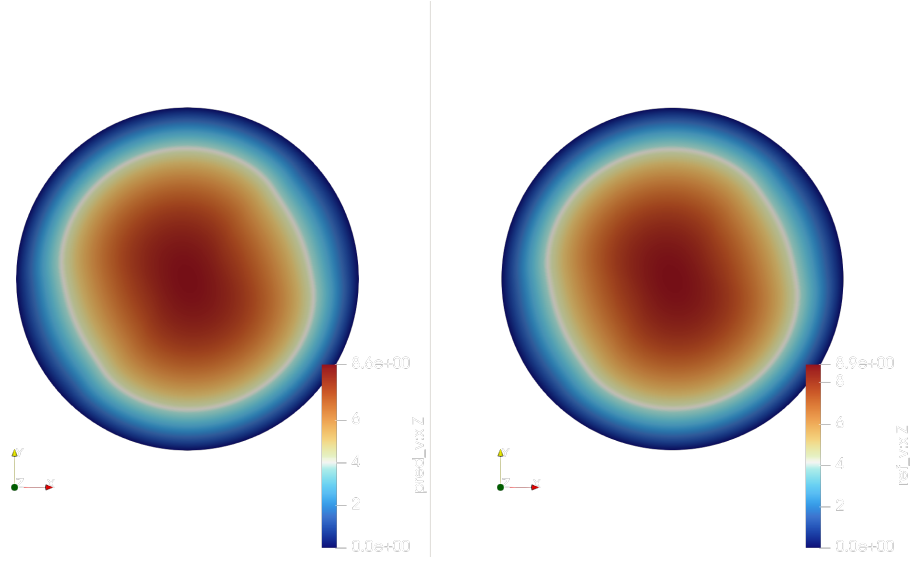


Рис. 1.8. Результаты численного эксперимента: поле деформаций при различных типах нагружения. Показаны компоненты деформации u_{11} , u_{22} и u_{12} для различных конфигураций нагружения.

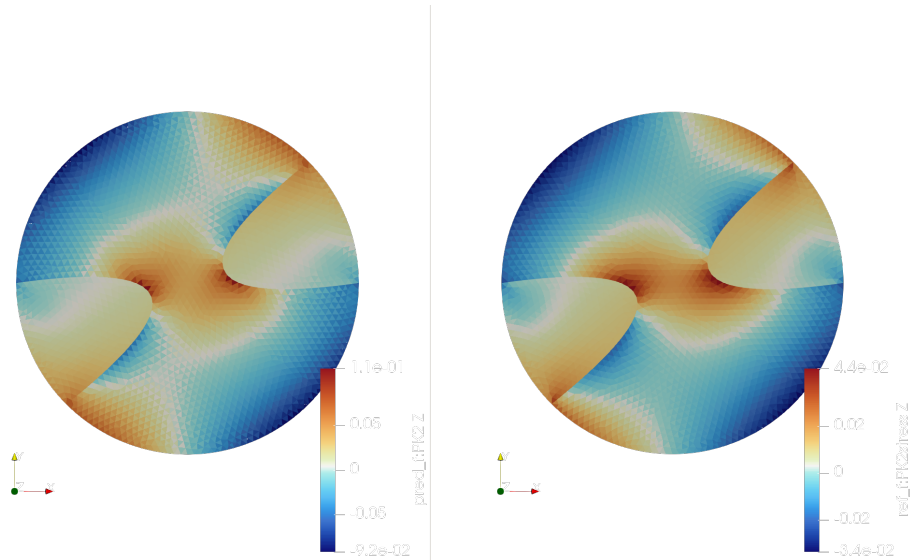


Рис. 1.9. Результаты численного эксперимента: поле напряжений ПК2. Показаны компоненты напряжений S_{11} , S_{22} и S_{12} , вычисленные моделью CLANN для соответствующих деформаций.

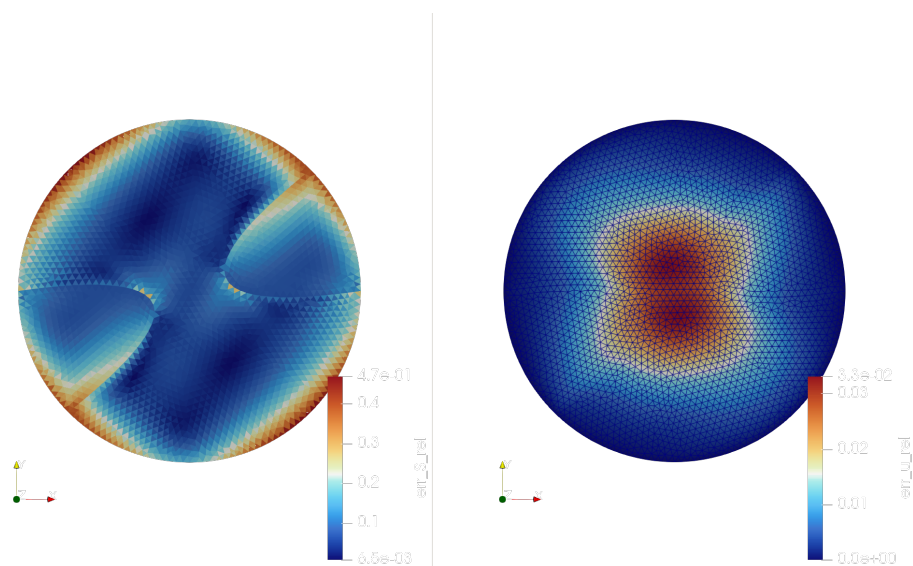


Рис. 1.10. Анализ ошибок численного эксперимента: распределение относительных ошибок между предсказанными и эталонными значениями напряжений. Показаны локальные и глобальные метрики точности модели.

Приложение А

Эквивалентность QR-факторизации \mathbf{F} и разложения Холецкого $\mathbf{C} = \mathbf{F}^\top \mathbf{F}$ для вычисления логарифмических координат $\boldsymbol{\xi}$

А.1 Постановка и обозначения

Рассматривается двумерная гиперупругая кинематика. Пусть:

- $\mathbf{F} \in \mathbb{R}^{2 \times 2}$ — градиент деформации, $\det \mathbf{F} > 0$,
- $\mathbf{C} = \mathbf{F}^\top \mathbf{F}$ — правый тензор Коши–Грина (симметричный положительно определённый, SPD),
- Холецкий: $\mathbf{C} = \mathbf{U}^\top \mathbf{U}$, где \mathbf{U} — верхнетреугольная и $\text{diag}(\mathbf{U}) > 0$,
- Логарифмические координаты: $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3) = (\ln u_{11}, \ln u_{22}, u_{12}/u_{11})$.

Цель: показать, что при наличии \mathbf{F} можно заменить вычисление $\mathbf{U} = \text{chol}(\mathbf{C})$ на $\mathbf{U} = \mathbf{R}$ из тонкого QR(\mathbf{F}) = $\mathbf{Q}\mathbf{R}$ (с $\text{diag}(\mathbf{R}) > 0$), и получить те же $\boldsymbol{\xi}$.

А.2 Теорема (эквивалентность \mathbf{U} и \mathbf{R})

Пусть $\mathbf{F} \in \mathbb{R}^{2 \times 2}$ невырождённая ($\det \mathbf{F} > 0$). Рассмотрим тонкую QR-факторизацию

$$\mathbf{F} = \mathbf{Q}\mathbf{R}, \quad (\text{A.1})$$

где $\mathbf{Q} \in \mathbb{R}^{2 \times 2}$ — ортогональная ($\mathbf{Q}^\top \mathbf{Q} = \mathbf{I}$), $\mathbf{R} \in \mathbb{R}^{2 \times 2}$ — верхнетреугольная. Выберем стандартную нормализацию $\text{diag}(\mathbf{R}) > 0$. Тогда \mathbf{R} совпадает с фактором Холецкого для \mathbf{C} :

$$\mathbf{R} = \text{chol}(\mathbf{C}), \quad \text{с} \quad \mathbf{C} = \mathbf{F}^\top \mathbf{F}. \quad (\text{A.2})$$

Доказательство.

$$\mathbf{C} = \mathbf{F}^\top \mathbf{F} = (\mathbf{QR})^\top (\mathbf{QR}) = \mathbf{R}^\top \mathbf{Q}^\top \mathbf{Q} \mathbf{R} = \mathbf{R}^\top \mathbf{R}. \quad (\text{A.3})$$

Так как \mathbf{C} — SPD и \mathbf{R} — верхнетреугольная с положительной диагональю, то представление $\mathbf{C} = \mathbf{R}^\top \mathbf{R}$ единственно. По единственности фактора Холецкого (с $\text{diag} > 0$) следует $\mathbf{R} = \text{chol}(\mathbf{C})$. \square

Следствие. Логарифмические координаты $\boldsymbol{\xi}$, определённые через $\mathbf{U} = \text{chol}(\mathbf{C})$, можно эквивалентно вычислять из $\mathbf{U} = \mathbf{R}$ в QR(\mathbf{F}), при условии $\text{diag}(\mathbf{R}) > 0$.

A.3 Координаты $\boldsymbol{\xi}$ через \mathbf{U}

$$\text{Для } \mathbf{U} = \begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix}, \text{diag}(\mathbf{U}) > 0,$$

$$\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3) = (\ln u_{11}, \ln u_{22}, u_{12}/u_{11}). \quad (\text{A.4})$$

Тем самым, $\boldsymbol{\xi}(\mathbf{F}) := \boldsymbol{\xi}(\mathbf{R}(\mathbf{F})) = \boldsymbol{\xi}(\mathbf{U}(\mathbf{C}))$.

Список литературы

- [1] Brandon Amos, Lei Xu, and J Zico Kolter. Input convex neural networks. In *International conference on machine learning*, pages 146–155. PMLR, 2017.
- [2] Stuart S Antman. *Nonlinear problems of elasticity*. Springer, 2005.
- [3] Charles Dugas, Yoshua Bengio, François Bélisle, Claude Nadeau, and René Garcia. Incorporating second-order functional knowledge for better option pricing. *Advances in neural information processing systems*, 13, 2001.
- [4] George Green. On the laws of the reflection and refraction of light at the common surface of two non-crystallized media. *Transactions of the Cambridge Philosophical Society*, 7:1–24, 1839.
- [5] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [6] Gustav Kirchhoff. Ueber das gleichgewicht und die bewegung einer elastischen scheibe. *Journal für die reine und angewandte Mathematik*, 40:51–88, 1850.
- [7] Alexey Liogky and Victoria Salamatova. Data-driven anisotropic biomembrane simulation based on the laplace stretch. *Computation*, 12(3):39, 2024.
- [8] R.W. Ogden. *Non-Linear Elastic Deformations*. Dover Publications, 1997.
- [9] Victoria Salamatova and Alexey Liogky. Interpretable data-driven modeling of hyperelastic membranes. *International Journal for Numerical Methods in Biomedical Engineering*, 39(11):e3757, 2023.
- [10] Clifford Truesdell. Historical introit the origins of rational thermodynamics. In *Rational thermodynamics*, pages 1–48. Springer, 1984.
- [11] Clifford Truesdell and Walter Noll. *The non-linear field theories of mechanics*. Springer Science & Business Media, 2004.