Оглавление

1	CLANN							
	1.1 Введение							
	1.2	2 Кинематика						
	1.3							
	1.4							
		1.4.1	Функция потерь и процесс обучения: физические ограничения					
			и регуляризация	6				
	1.5	5 Виртуальный эксперимент						
		1.5.1	Интерполяция и экстраполяция кривых нагружения	10				
		1.5.2	Виртуальные эксперименты	10				
	1.6	3 Заключение						
${f A}$	Эквивалентность QR-факторизации F и разложения Холецкого							
$\mathbf{C} {=} \mathbf{F}^{ op} \mathbf{F}$ для вычисления логарифмических координат $oldsymbol{\xi}$								
	новка и обозначения	12						
А.2 Теорема (эквивалентность U и R)								
	A.3	Коорд	инаты $m{\mathcal{E}}$ через $m{U}$	1.9				

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 б проверено при помощи обучения модели на синтетических и натурных 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 срединной поверхности. Обозначим через \mathbf{X} и \mathbf{x} положения точек в исходной (недеформированной) Ω_0 и текущей (деформированной) Ω_t конфигурациях срединной поверхности мембраны соответственно. Деформация определяется отображением $\mathbf{x} = \mathbf{x}(\mathbf{X})$, поверхностный градиент деформации $\mathbf{F} = \frac{\partial \mathbf{x}}{\partial \mathbf{X}}$, а правый тензор Коши—Грина $\mathbf{C} = \mathbf{F}^{\top}\mathbf{F}$. Для определения меры деформации мы используем тензор Лапласа $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3)$ [5], который 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 \mathbf{A}). В этом случае гиперупругий потенциал является функцией от деформации Лапласа $\psi = \psi(\boldsymbol{\xi})$.

Мера деформации Лапласа В двумерном случае вводятся координаты

$$\xi_1 = \ln(u_{11}), \quad \xi_2 = \ln(u_{22}), \quad \xi_3 = \frac{u_{12}}{u_{11}}, \quad u_{ij}$$
 – компоненты тензора \boldsymbol{U} . (1.1)

1.3 Напряжение и термодинамическая корректность

Гиперупругие напряжения вычисляется по цепному правилу дифференцированием энергии ψ по правому тензору деформации Коши-Грина C:

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

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} инвариантна к поворотам.
- Симметрия напряжений: $\mathbf{S} = \mathbf{S}^{\top}$ вследствие симметрии \mathbf{C} и корректного применения цепного правила.
- **Термодинамическая корректность:** равенство (1.2) обеспечивает согласованность с вторым законом термодинамики.

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:

$$S_{11} = e^{-2\xi_1} (r_1 - 2\xi_3 r_3) + e^{-2\xi_2} r_2 \xi_3^2,$$

$$S_{22} = e^{-2\xi_2} r_2,$$

$$S_{12} = -e^{-2\xi_2} r_2 \xi_3 + e^{-2\xi_1} r_3.$$
(1.3)

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:

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

Эти свойства принято записывать через градиент деформации и правый тензор деформа Коши-Грина [1], но они эквавалетны и для ξ .

Напряжения в естесственной конфигурации

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

$$S \to \mathbf{0}$$
 при $\|\xi\| \to 0$, (1.5)

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

1.4 Архитектура CLaNN и её производные

В рамках предложенного подхода CLaNN (Convex Laplace Neural Network) энергия деформации $\psi(\xi)$ с мерой деформации Лапласа аппроксимириуется посредством выпуклой по входу нейронной сетью (Input Convex Neural Network, ICNN) и вычисления 2 тензора напряжения Пиолы-Кирхгофа S.

Обобщенная архитектура ICNN

ICNN представляет собой класс нейронных сетей, гарантирующих выпуклость выходной функции относительно входных переменных. Общая архитектура ICNN с L скрытыми слоями определяется следующими соотношениями:

$$z_1 = \sigma(\mathbf{W}_1 \boldsymbol{\xi} + \mathbf{b}_1), \tag{1.6}$$

$$z_{i+1} = \sigma(\mathbf{W}_{i+1}z_i + \mathbf{U}_{i+1}\boldsymbol{\xi} + \mathbf{b}_{i+1}), \quad i = 1, \dots, L-1,$$
 (1.7)

$$\psi = \mathbf{W}_{L+1} z_L + \mathbf{U}_{L+1} \xi + b_{L+1}, \tag{1.8}$$

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 и одним выходным нейроном, определяемая следующими соотношениями:

$$s = \mathbf{W}_1 \boldsymbol{\xi} + \mathbf{b}_1, \qquad z = \frac{\text{softplus}(\beta s)}{\beta}, \qquad \psi = \mathbf{W}_2^{\mathsf{T}} z + b_2, \qquad \mathbf{W}_2 \ge 0.$$
 (1.9)

Функция активации softplus $(\beta x)/\beta$ строго выпукла и $\lim_{\beta\to\infty}$ softplus $(\beta x)/\beta=\max(0,x)$ (ReLU). Условие $\mathbf{W}_2\geq 0$ сохраняет выпуклость линейной комбинации.

После прохождения через все ребра вычислительного графа и получения ψ происходит автоматическое дифференцирование $\partial \psi/\partial \boldsymbol{\xi}$, которое реализована во всех современных бибилотеках для машинного обучения, после чего вычисляется тензор

напряжений S по формуле (1.3).

Размерности весовых параметров сети: $\mathbf{W}_1 \in \mathbb{R}^{h \times 3}$, $\mathbf{b}_1 \in \mathbb{R}^h$, $\mathbf{W}_2 \in \mathbb{R}^h_{\geq 0}$, $b_2 \in \mathbb{R}$, где h — размерность скрытого слоя.

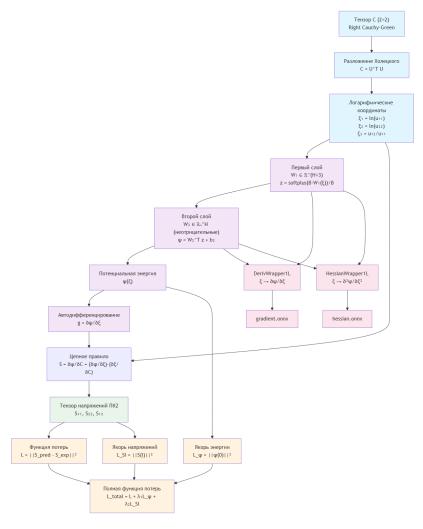


Рис. 1.1. Схема вычислительного процесса CLANN: от входного тензора до функции потерь. Показаны этапы обработки входных данных, вычисления нейросетью, дифференцирования и формирования функции потерь.

Аналитические выражения для производных энергии Градиент энергии деформации

Аналитическое дифференцирование функции энергии по переменным ξ даёт выражение для градиента:

$$\nabla_{\boldsymbol{\xi}} \psi = \boldsymbol{W}_{1}^{T} \left(\boldsymbol{W}_{2} \odot \sigma(\beta(\boldsymbol{W}_{1}\boldsymbol{\xi} + \boldsymbol{b}_{1})) \right), \tag{1.10}$$

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

Гессиан энергии деформации

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

$$H_{ij} = \sum_{h} \sigma'_{h} w_{2,h} W_{h,i} W_{h,j}, \qquad (1.11)$$

где $\sigma'=\beta\,\sigma(1-\sigma)$ - производная сигмоидальной функции, $\sigma=\mathrm{sigmoid}(\beta s)$, а $s=\pmb{W}_1\xi+\pmb{b}_1.$

Материальная устойчивость и положительная определённость

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

$$\boldsymbol{H} = \frac{\partial^2 \psi}{\partial \xi^2} > 0, \tag{1.12}$$

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

1.4.1 Функция потерь и процесс обучения: физические ограничения и регуляризация

Основная функция потерь: минимизация невязки напряжений

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

$$L = \frac{1}{N} \sum_{i=1}^{N} ||\mathbf{S}_{\text{pred}}^{(i)} - \mathbf{S}_{\text{exp}}^{(i)}||^{2}.$$
 (1.13)

Физические ограничения и регуляризация

Для обеспечения физической корректности модели в функцию потерь вводятся дополнительные слагаемые, отражающие принципы механики:

$$L_{\text{SI}} = \|\mathbf{S}(\mathbf{I})\|^2, \qquad L_{\psi} = \|\psi(0)\|^2.$$
 (1.14)

Первое слагаемое $L_{\rm SI}$ обеспечивает выполнение условия естественного состояния: при отсутствии деформации (C=I) тензор напряжений должен обращаться в нуль. Это условие является прямым следствием принципа объективности и соответствует физическому требованию отсутствия внутренних напряжений в недеформированном состоянии.

Второе слагаемое L_{ψ} гарантирует, что энергия деформации в естественном состоянии равна нулю.

Полная функция потерь и процесс оптимизации

Объединение основного функционала потерь с физическими ограничениями даёт полную функцию потерь:

$$L_{\text{total}} = L + \lambda_{\text{SI}} L_{\text{SI}} + \lambda_{\psi} L_{\psi}. \tag{1.15}$$

Коэффициенты λ_{SI} и λ_{ψ} определяют относительную важность физических ограничений по сравнению с точностью аппроксимации экспериментальных данных. Их выбор является критически важным для баланса между точностью модели и физической корректностью.

Для минимизации полной функции потерь (1.15) используется оптимизатор Adam [2], 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 коррект**ность** достигается через строгое соблюдение соотношения (1.2), что гарантирует консервативность напряжений ϕS : $\mathrm{d} C = 0$ и согласованность с законами термодинамики; материальная устойчивость обеспечивается строгой выпуклостью функции энергии $\psi(\boldsymbol{\xi})$, гарантируемой архитектурой ICNN с неотрицательными весами $W_2 \ge 0$ и выпуклой функцией активации, что влечёт положительную определённость гессиана (1.11) и единственность решения; объективность автоматически выполняется благодаря параметризации через тензор Коши-Грина $\boldsymbol{C} = \boldsymbol{F}^{\top} \boldsymbol{F}$, обеспечивая инвариантность относительно поворотов и симметрию напряжений; положительность энергии гарантируется комбинацией неотрицательных весов и положительной функции активации, что обеспечивает $\psi(\boldsymbol{\xi}) > 0$ и $\psi(\boldsymbol{\xi}) \to \infty$ при $\|\boldsymbol{\xi}\| \to \infty$; численная стабильность достигается через логарифмическую параметризацию Лапласа (1.1), которая корректно обрабатывает большие деформации и предотвращает нефизичные значения; наконец, физические ограничения обеспечиваются через якорные условия в функции потерь (1.14), гарантирующие выполнение условий естественного состояния S(I) = 0 и $\psi(0) = 0$.

1.5 Виртуальный эксперимент

Мы используем синтетические экспериментальные данные для тестирования CLaNN, для изотропного надувания 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 [4]

$$\psi = \frac{\mu}{2}(I_1 - 3),\tag{1.16}$$

где $\mu=0.432\,\mathrm{M\Pi a}$ - модуль сдвига, а $I_1=\mathrm{tr}(\boldsymbol{C})$ - первый инвариант правого тензора деформаций Коши-Грина \boldsymbol{C} .

Причем данные для обучения собирались из одного центрального элемента сетки, что соответствует ограничениям эквивалетного 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 [3].

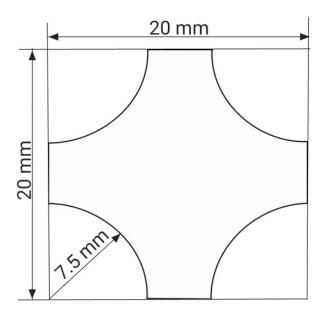


Рис. 1.2. Размеры образца биоматериала в форме мальтийского креста. Радиус вырезов одинаков для всех вырезов

Схематическое представление протоколов показано на рисунке, где $w_i \in [0,1]$, $i \in \{1,4\}$ представляет собой процент от заданного максимального смещения u_{\max} для i-го плеча: $w_i = 0$ соответствует закрепленному плечу, а $w_i = 1$ — максимальному смещению. 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 $n = 1, \ldots, N$, $N = u_{\max}/\Delta s$ — количество шагов. Треугольная сетка для образца является квазиравномерной с размером ячейки $h_{\mathrm{fit}} = 0.25$ мм, максимальное смещение $u_{\max} = 2$ мм и $\Delta s = 0.025$ мм. На каждом шаге мы собираем данные (C, S) для всех треугольников, принадлежащих центральной области. Поскольку мы используем линейные (P_1) конечные элементы, значения (C, S) постоянны на каждом треугольнике.

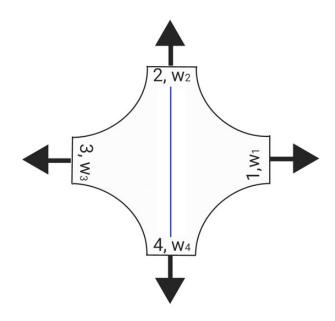


Рис. 1.3. Схематическое представление протоколов. Радиус вырезов одинаков для всех вырезов

Наш предлагаемый тестовый протокол предполагает девять экспериментов:

Таблица 1.1. Протоколы тестовых экспериментов

Nº	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	1	1/3	1	1/3
6	0.75	1	0.75	1
7	1	0.75	1	0.75
8	0.5	1	0.5	1
9	1/3	1	1/3	1

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

Гиперпараметры оптимизации:

- Скорость обучения (learning rate): 0.001
- Размер батча (batch size): 128
- Веса физических ограничений: $\lambda_{\rm SI}=0.1,\,\lambda_{\psi}=0.1$
- Архитектура: 16 нейронов на скрытом слое

Результаты обучения: Процесс оптимизации показал высокую эффективность: ошибка аппроксимации снизилась на 5 порядков за менее чем 5000 эпох, 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.

1.5.1 Интерполяция и экстраполяция кривых нагружения

вставить рисунок

1.5.2 Виртуальные эксперименты

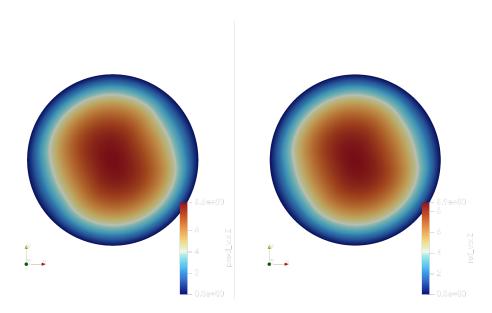


Рис. 1.4. Результаты численного эксперимента: поле деформаций при различных типах нагружения. Показаны компоненты деформации u_{11} , u_{22} и u_{12} для различных конфигураций нагружения.

1.6 Заключение

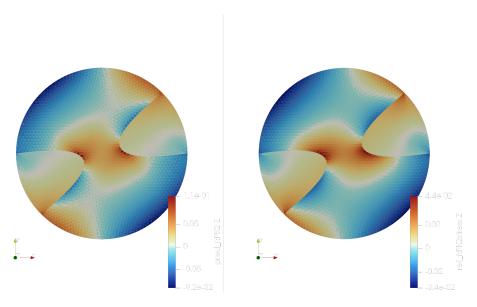


Рис. 1.5. Результаты численного эксперимента: поле напряжений ПК2. Показаны компоненты напряжений S_{11} , S_{22} и S_{12} , вычисленные моделью CLANN для соответствующих деформаций.

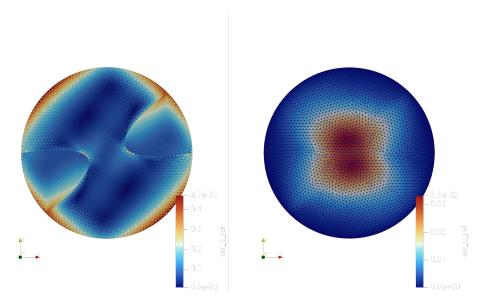


Рис. 1.6. Анализ ошибок численного эксперимента: распределение относительных ошибок между предсказанными и эталонными значениями напряжений. Показаны локальные и глобальные метрики точности модели.

Приложение А

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

А.1 Постановка и обозначения

Рассматривается двумерная гиперупругая кинематика. Пусть:

- $\boldsymbol{F} \in \mathbb{R}^{2 \times 2}$ градиент деформации, $\det \boldsymbol{F} > 0$,
- $C = F^{\top}F$ правый тензор Коши-Грина (симметричный положительно определённый, SPD),
- ullet Холецкий: $oldsymbol{C} = oldsymbol{U}^{ op} oldsymbol{U}$, где $oldsymbol{U}$ верхнетреугольная и $\mathrm{diag}(oldsymbol{U}) > 0$,
- 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 \pmb{F} можно заменить вычисление $\pmb{U}=\operatorname{chol}(\pmb{C})$ на $\pmb{U}=\pmb{R}$ из тонкого $\operatorname{QR}(\pmb{F})=\pmb{Q}\pmb{R}$ (c $\operatorname{diag}(\pmb{R})>0$), и получить те же $\pmb{\xi}$.

А.2 Теорема (эквивалентность U и R)

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

$$\boldsymbol{F} = \boldsymbol{Q}\boldsymbol{R},\tag{A.1}$$

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

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

Доказательство.

$$C = F^{\mathsf{T}} F = (QR)^{\mathsf{T}} (QR) = R^{\mathsf{T}} Q^{\mathsf{T}} QR = R^{\mathsf{T}} R. \tag{A.3}$$

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

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

${ m A.3}$ Координаты $oldsymbol{\xi}$ через $oldsymbol{U}$

Для
$$\boldsymbol{U} = \begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix}$$
, $\operatorname{diag}(\boldsymbol{U}) > 0$,
$$\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3) = (\ln u_{11}, \ln u_{22}, u_{12}/u_{11}). \tag{A.4}$$

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

Список литературы

- [1] Stuart S Antman. Nonlinear problems of elasticity. Springer, 2005.
- [2] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.
- [3] Alexey Liogky and Victoria Salamatova. Data-driven anisotropic biomembrane simulation based on the laplace stretch. *Computation*, 12(3):39, 2024.
- [4] R.W. Ogden. Non-Linear Elastic Deformations. Dover Publications, 1997.
- [5] Victoria Salamatova and Alexey Liogky. Interpretable data-driven modeling of hyperelastic membranes. *International Journal for Numerical Methods in Biomedical Engineering*, 39(11):e3757, 2023.