

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ  
МОСКОВСКИЙ ФИЗИКО-ТЕХНИЧЕСКИЙ ИНСТИТУТ (государственный университет)  
ФИЗТЕХ-ШКОЛА ПРИКЛАДНОЙ МАТЕМАТИКИ И ИНФОРМАТИКИ  
КАФЕДРА «ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ»

Карпев Глеб Андреевич

**Порождающие модели для прогнозирования наборов  
временных рядов**

010990 — Интеллектуальный анализ данных

Выпускная квалификационная работа бакалавра

**Научный руководитель:**  
д. ф.-м. н. Стрижов Вадим Викторович

Москва  
2025

# Содержание

|   |           |
|---|-----------|
| <b>1 Введение</b>   | <b>4</b>  |
| <b>2 Основные обозначения и сокращения</b>                            | <b>4</b>  |
| <b>3 Постановка задачи</b>  | <b>5</b>  |
| <b>4 Обзор литературы</b>   | <b>6</b>  |
| 4.1 Генеративные модели на основе оценки градиента (SGM) . . . . .    | 6         |
| 4.2 Геометрия многообразия SPD-матриц . . . . .                       | 8         |
| 4.2.1 Аффинно-инвариантная метрика и расстояние . . . . .             | 8         |
| 4.2.2 Касательные пространства и отображения . . . . .                | 8         |
| 4.3 Римановы генеративные модели (RSGM) . . . . .                     | 9         |
| 4.4 Римановы генеративные модели (RSGM) . . . . .                     | 11        |
| <b>5 Предложенный метод</b>   | <b>12</b> |
| 5.1 Условная генерация ковариационных матриц с помощью RSGM . . . . . | 12        |
| 5.2 Процесс обучения и прогнозирования . . . . .                      | 13        |
| <b>6 Вычислительный эксперимент</b>                                   | <b>14</b> |
| 6.1 Экспериментальные данные . . . . .                                | 14        |
| <b>7 Заключение</b>   | <b>14</b> |
| <b>Список литературы</b>  | <b>14</b> |

## Аннотация

Решается задача прогнозирования наборов временных рядов с высокой ковариацией и высокой дисперсией. Для решения данной задачи предлагается построить пространство парных расстояний, представляющего метрическую конфигурацию временных рядов. Прогноз осуществляется в данном пространстве, а затем результат возвращается в исходное пространство с использованием метода многомерного шкалирования. В данной работе предлагаются порождающие модели для прогнозирования наборов временных рядов в метрическом вероятностном пространстве. Новизна работы заключается в применении римановых моделей для прогнозирования и использовании римановых генеративных диффузных моделей.

**Ключевые слова:** *многообразие SPD-матриц, Riemannian Score-based Generative Models, conditional generation*

# 1 Введение

В данной работе предлагается метод прогнозирования набора временных рядов с высокой ковариацией и высокой дисперсией. Предлагаемое решение задачи прогнозирования состоит из трех этапов. Во-первых, построение пространства парных расстояний, где используемая метрика удовлетворяет условию Мерсера. Во-вторых, прогнозируется матрица попарных расстояний. В-третьих, результат возвращается в исходное пространство. В данной работе изучаются римановы генеративные диффузные модели (RSGMs) [1] для выполнения прогнозирования матрицы попарных расстояний.

Классическими алгоритмами прогнозирования временных рядов являются метод SSA (Singular Spectrum Analysis) [2], рекуррентные нейронные сети с долгой краткосрочной памятью (LSTM) [3], State space model [4]. Новизна предложенного метода заключается в том, что выполняется кодирование временных рядов с помощью матрицы расстояний, выполняется прогноз, а затем декодирование полученной матрицы.

Генеративные диффузионные модели, также называемые Score-based generative models (SGMs), представляют собой мощный класс порождающих моделей [5, 6]. SGM состоит из этапа зашумления, на котором к данным постепенно добавляется гауссовский шум, и процесса шумоподавления, на котором генеративная модель постепенно снижает гауссовский шум, чтобы получить образец данных. Существующие SGM предполагают, что данные находятся в евклидовом пространстве, но зачастую данные естественным образом описываются с помощью римановых многообразий, и методы SGM не подходят. Предлагается использовать римановы генеративные диффузные модели (RSGMs) [1], класс генеративных моделей, расширяющих SGM [5, 6] до римановых многообразий.

Анализ предлагаемого метода прогнозирования проводится на синтетических и реальных данных. Синтетический набор данных построен на основе синусоидальных сигналах со случайной амплитудой и частотой. Реальные данные представляют собой временной ряд цены на электроэнергию. Цель эксперимента — показать работоспособность предложенного подхода и провести его сравнение с базовым алгоритмом. В качестве базовой модели для сравнения используется модель, в которой прогнозное значение ковариационной матрицы на следующем шаге полагается равным ее значению на предыдущем шаге. Получена оценка качества работы построенного алгоритма.

Данная работа имеет следующую структуру. Раздел 2 вводит основные обозначения. В Разделе 3 приводится формальная постановка задачи. Раздел 4 представляет обзор по генеративным моделям и основам римановой геометрии. В Разделе 5 детально описывается предложенный метод с использованием условной генерации. Раздел 6 содержит описание и результаты вычислительных экспериментов. В Разделе 7 подводятся итоги работы и обсуждаются дальнейшие направления исследований.

## 2 Основные обозначения и сокращения

В тексте настоящей работы приняты следующие математические обозначения. Скалярные величины обозначаются строчными буквами курсивом, например,  $t$  или  $n$ .

Векторы представляются строчными полужирными буквами, например,  $\mathbf{x}_t$ . Матрицы обозначаются прописными полужирными буквами, например,  $\mathbf{C}_t$ . Операция транспонирования вектора или матрицы обозначается надстрочным индексом  $\mathbf{T}$ .

Помимо математических символов, в работе используется ряд сокращений для обозначения ключевых понятий и методов. Сокращение SGM (Score-based Generative Models) используется для обозначения порождающих моделей на основе оценки градиента. Их обобщение на случай неевклидовых пространств обозначается как RSGM (Riemannian Score-based Generative Models).

Ключевым объектом в предлагаемом методе являются симметричные положительно определенные матрицы, для которых используется аббревиатура SPD (Symmetric Positive-Definite).

Для обозначения методов спектрального анализа временных рядов, упоминаемых в обзоре литературы, используются сокращения SSA (Singular Spectrum Analysis) и MSSA (Multivariate Singular Spectrum Analysis) для одномерного и многомерного случаев соответственно.

### 3 Постановка задачи

Рассматривается задача прогнозирования набора из  $n$  временных рядов. В каждый дискретный момент времени  $t$  состояние системы представлено в виде вектора-столбца  $\mathbf{x}_t = [x_{1t}, \dots, x_{nt}]^T$ . Предполагается, что временные ряды в наборе характеризуются высокой степенью взаимной корреляции и значительной дисперсией, что усложняет их совместное моделирование.

Основная идея предлагаемого подхода заключается в переходе от прямого прогнозирования векторов  $\mathbf{x}_t$  к прогнозированию их совместной статистической структуры. Для этого в каждый момент времени  $t$  совокупность исторических данных о наборе временных рядов ставится в соответствие с матрицей ковариаций  $\mathbf{C}_t$ , которая вычисляется по следующей формуле:

$$\mathbf{C}_t = \frac{1}{t-1} \mathbf{Y}_t \mathbf{Y}_t^T, \quad (1)$$

где  $\mathbf{Y}_t$  — это матрица, столбцы которой представляют собой центрированные векторы временных рядов:

$$\mathbf{Y}_t = [\mathbf{x}_1 - \mu_t, \dots, \mathbf{x}_t - \mu_t]. \quad (2)$$

Здесь  $\mu_t$  является вектором среднего значения, вычисленным по всем доступным историческим данным до момента  $t$ :

$$\mu_t = \frac{1}{t} \sum_{i=1}^t \mathbf{x}_i. \quad (3)$$

Таким образом, задача сводится к нахождению оптимального отображения  $f$ , действующего в пространстве симметричных положительно определенных матриц размерности  $n \times n$ , которое обозначается как  $Sym^+(n)$ . Это отображение

$$f : Sym^+(n) \rightarrow Sym^+(n)$$

должно прогнозировать будущую ковариационную матрицу  $\mathbf{C}_{t+\Delta}$  на основе текущей  $\mathbf{C}_t$ . Качество прогноза  $\hat{\mathbf{C}}_{t+\Delta} = f(\mathbf{C}_t)$  оценивается с помощью метрики, учитывающей

внутреннюю геометрию пространства  $Sym^+(n)$ . В качестве функции потерь используется ожидание квадрата геодезического расстояния между истинной и спрогнозированной матрицами. Задача оптимизации формулируется как поиск такого отображения  $f$ , которое минимизирует следующую функцию ошибки:

$$\mathcal{L} = \mathbb{E} [d_{\text{Riem}}^2 (\mathbf{C}_{t+\Delta}^{\text{true}}, f(\mathbf{C}_t))] \rightarrow \min_f, \quad (4)$$

где  $d_{\text{Riem}}(\cdot, \cdot)$  обозначает риманово геодезическое расстояние на многообразии  $Sym^+(n)$ .

## 4 Обзор литературы

В данной главе рассматриваются теоретические основы, составляющие фундамент предложенного в работе метода. Сначала будет дан обзор генеративных моделей на основе оценки градиента (SGM), затем будут введены ключевые понятия из римановой геометрии, касающиеся многообразия симметричных положительно определенных (SPD) матриц, и, наконец, будет описано их обобщение — римановы генеративные модели (RSGM).

### 4.1 Генеративные модели на основе оценки градиента (SGM)

Генеративные модели на основе оценки градиента (Score-based Generative Models, SGM), также известные как диффузионные модели, представляют собой мощный класс порождающих моделей, продемонстрировавших передовые результаты в различных областях [5, 7]. Методология SGM состоит из двух ключевых процессов, которые описываются стохастическими дифференциальными уравнениями (SDE) [8].

**Прямой процесс (зашумление).** Этот процесс постепенно преобразует сложное распределение данных  $p_0(\mathbf{x})$  в простое, априорное распределение  $p_T(\mathbf{x})$ , как правило, в стандартное гауссовское  $\mathcal{N}(0, \mathbf{I})$ . Процесс описывается SDE общего вида:

$$d\mathbf{x} = f(\mathbf{x}, t)dt + g(t)d\mathbf{W}, \quad (5)$$

где  $t \in [0, T]$ ,  $f(\mathbf{x}, t)$  — векторная функция дрейфа (drift),  $g(t)$  — скалярная функция диффузии (diffusion), а  $\mathbf{W}$  — стандартный винеровский процесс. Функции  $f$  и  $g$  выбираются таким образом, чтобы при  $t \rightarrow T$  распределение  $p_t(\mathbf{x})$  сходилось к априорному, а переходное распределение  $p_{t|s}(\mathbf{x}_t | \mathbf{x}_s)$  для  $s < t$  было гауссовским и имело аналитическое выражение.

**Обратный процесс (генерация).** Согласно теореме Андерсона [9], существует обратное по времени SDE, которое описывает процесс, обращающий диффузию, заданную уравнением (5). Это обратное SDE имеет вид:

$$d\mathbf{x} = [f(\mathbf{x}, t) - g(t)^2 \nabla_{\mathbf{x}} \log p_t(\mathbf{x})] dt + g(t)d\bar{\mathbf{W}}, \quad (6)$$

где  $\bar{\mathbf{W}}$  — винеровский процесс при обратном течении времени, а  $\nabla_{\mathbf{x}} \log p_t(\mathbf{x})$  — это градиент логарифма плотности данных в момент времени  $t$ . Этот градиент называется *функцией оценки* (score function).

Решение этого обратного уравнения, начиная с сэмпла  $\mathbf{x}_T \sim p_T$ , позволяет сгенерировать новый объект из исходного распределения данных  $\mathbf{x}_0 \sim p_0$ .

**Аппроксимация функции оценки.** Ключевая сложность заключается в том, что функция оценки  $\nabla_{\mathbf{x}} \log p_t(\mathbf{x})$  неизвестна, так как распределение  $p_t(\mathbf{x})$  недоступно в явном виде. Для ее аппроксимации используется нейронная сеть  $s_{\theta}(\mathbf{x}, t)$ , обучаемая с помощью метода, известного как Denoising Score Matching [10]. Целевая функция для обучения нейросети имеет вид:

$$\mathcal{L}(\theta) = \mathbb{E}_{t, \mathbf{x}_0, \mathbf{x}_t} \left[ \lambda(t) \left\| s_{\theta}(\mathbf{x}_t, t) - \nabla_{\mathbf{x}_t} \log p_{t|0}(\mathbf{x}_t | \mathbf{x}_0) \right\|^2 \right], \quad (7)$$

где математическое ожидание берется по времени  $t \in [0, T]$ , сэмплам из исходного распределения  $\mathbf{x}_0 \sim p_0(\mathbf{x})$  и сэмплам  $\mathbf{x}_t \sim p_{t|0}(\mathbf{x}_t | \mathbf{x}_0)$ . Преимущество этого подхода заключается в том, что условная плотность  $p_{t|0}(\mathbf{x}_t | \mathbf{x}_0)$  известна аналитически (является гауссовской), и ее градиент легко вычисляется. Функция  $\lambda(t)$  является весовой функцией, которая влияет на качество генерации.

Основным ограничением стандартных SGM является то, что все операции, включая вычисление градиента, определены в евклидовом пространстве  $\mathbb{R}^n$ . Это делает их неприменимыми для данных, которые естественным образом существуют на многообразиях с нетривиальной геометрией. Общая схема работы SGM представлена на рисунке 1.

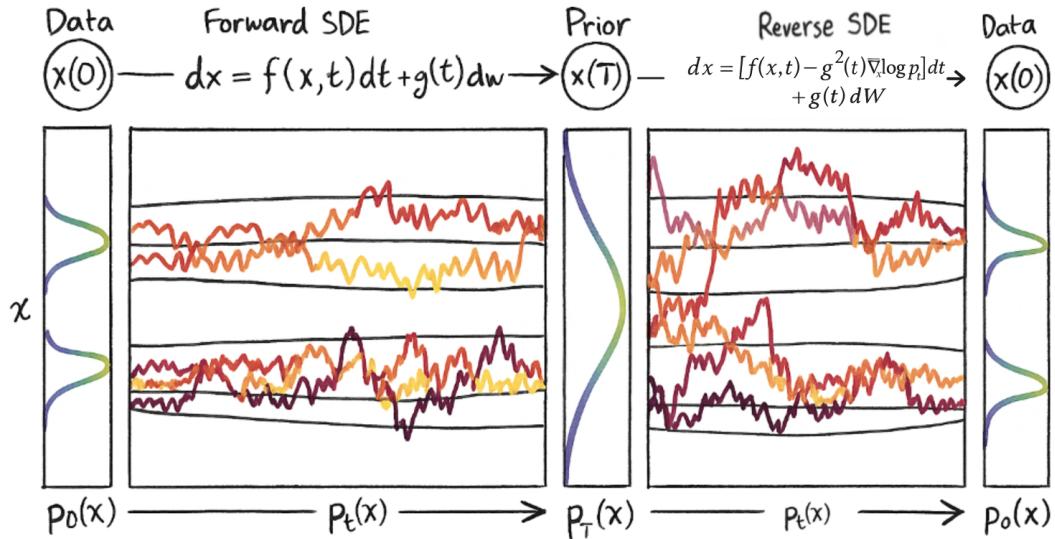


Рис. 1: Схематическое представление работы Score-based Generative Model (SGM). Прямой процесс постепенно добавляет шум к исходным данным  $p_0(\mathbf{x})$ , превращая их в простое гауссовское распределение  $p_T(\mathbf{x})$ . Обратный процесс начинается со случайног шума и, следуя оценкам градиента  $\nabla \log p_t$ , выученным моделью  $s_{\theta}$ , восстанавливает объект из исходного распределения данных.

## 4.2 Геометрия многообразия SPD-матриц

Ковариационные матрицы, являющиеся центральным объектом в данной работе, по определению являются симметричными и положительно определенными (Symmetric Positive-Definite, SPD). Множество всех SPD-матриц размера  $n \times n$  образует риманово многообразие. Далее в тексте мы будем обозначать это многообразие как  $\mathcal{M} = Sym^+(n)$  или, следуя обозначениям на рисунке 2, как  $\mathcal{S}_{++}$ . Это многообразие обладает богатой геометрической структурой, которую необходимо учитывать для корректного анализа и моделирования данных [11, 12]. Ключевое отличие заключается в том, что кратчайшим путем между двумя точками (матрицами) является кривая, называемая геодезической.

### 4.2.1 Аффинно-инвариантная метрика и расстояние

Центральным элементом риманова многообразия является метрический тензор, который определяет способ измерения расстояний и углов. Для многообразия  $\mathcal{M}$  стандартным выбором является аффинно-инвариантная метрика. Она обладает важным свойством инвариантности относительно аффинных преобразований исходных данных, что делает ее особенно подходящей для ковариационных матриц. Скалярное произведение двух векторов  $\mathbf{A}, \mathbf{B}$ , принадлежащих касательному пространству в точке  $\mathbf{G} \in \mathcal{M}$ , определяется как:

$$\langle \mathbf{A}, \mathbf{B} \rangle_{\mathbf{G}} = \text{tr}(\mathbf{G}^{-1} \mathbf{A} \mathbf{G}^{-1} \mathbf{B}). \quad (8)$$

На основе этой метрики определяется геодезическое расстояние, обозначаемое на рисунке 2 как  $\delta(\cdot, \cdot)$ . Расстояние между матрицами  $\mathbf{C}_1$  и  $\mathbf{G}$  вычисляется по формуле:

$$\delta_{\text{Riem}}(\mathbf{C}_1, \mathbf{G}) = \|\text{Log}_{\mathbf{G}}(\mathbf{C}_1)\|_F = \|\log(\mathbf{G}^{-1/2} \mathbf{C}_1 \mathbf{G}^{-1/2})\|_F, \quad (9)$$

где  $\text{Log}_{\mathbf{G}}(\mathbf{C}_1)$  — логарифмическое отображение,  $\log(\cdot)$  — операция взятия главного матричного логарифма, а  $\|\cdot\|_F$  — норма Фробениуса.

### 4.2.2 Касательные пространства и отображения

Для выполнения вычислений на многообразии необходимо уметь переходить от его криволинейной структуры к локальным приближениям. Эти пространства называются касательными пространствами. Касательное пространство  $T_{\mathbf{G}}\mathcal{M}$  в точке  $\mathbf{G}$  можно интуитивно представить как плоскость, касающуюся многообразия в этой точке. Векторы в этом пространстве представляют собой симметричные матрицы, обозначаемые на рисунке 2 как  $\mathcal{S}$ . Переход между многообразием ( $\mathcal{S}_{++}$ ) и его касательными пространствами ( $\mathcal{S}$ ) осуществляется с помощью двух ключевых взаимно обратных операций, которые схематически показаны на рисунке 2.

**Логарифмическое отображение (Log).** Данное отображение переносит точку  $\mathbf{C}$  с многообразия в касательное пространство в точке  $\mathbf{G}$ , представляя ее в виде вектора  $\xi$ , который указывает направление и «длину» геодезического пути от  $\mathbf{G}$  до  $\mathbf{C}$ . Формула логарифмического отображения имеет вид:

$$\xi = \text{Log}_{\mathbf{G}}(\mathbf{C}) = \mathbf{G}^{1/2} \log(\mathbf{G}^{-1/2} \mathbf{C} \mathbf{G}^{-1/2}) \mathbf{G}^{1/2}. \quad (10)$$

**Экспоненциальное отображение (Exp).** Данное отображение, обратное логарифмическому, «перемещает» точку  $\mathbf{G}$  вдоль геодезической в направлении, заданном вектором  $\xi$  из касательного пространства  $T_{\mathbf{G}}\mathcal{M}$ . Результатом является новая точка на многообразии. Формула экспоненциального отображения:

$$\mathbf{C} = \text{Exp}_{\mathbf{G}}(\xi) = \mathbf{G}^{1/2} \exp \left( \mathbf{G}^{-1/2} \xi \mathbf{G}^{-1/2} \right) \mathbf{G}^{1/2}, \quad (11)$$

где  $\exp(\cdot)$  — операция взятия матричной экспоненты. Эти два отображения являются основой для всех вычислений в предлагаемом методе, позволяя корректно работать с геометрией пространства ковариационных матриц.

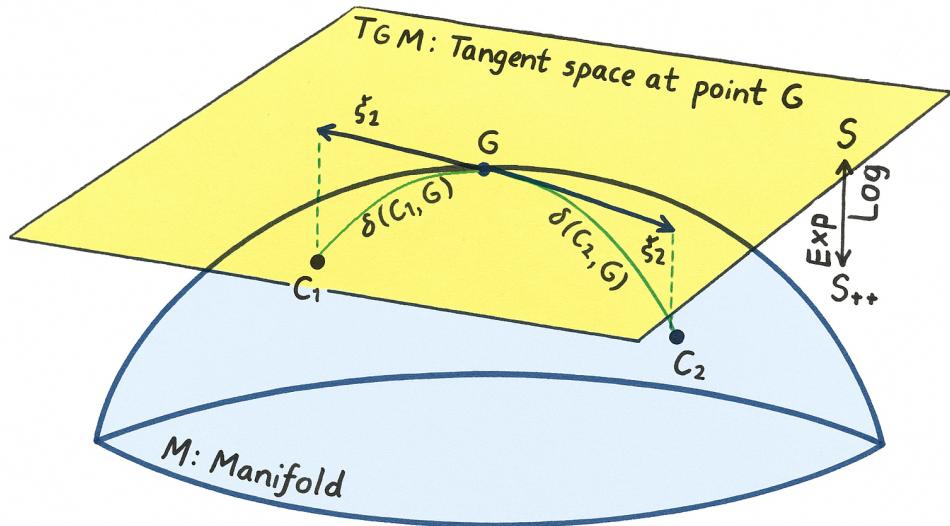


Рис. 2: Концептуальная иллюстрация риманова многообразия  $\mathcal{M}$  (также  $\mathcal{S}_{++}$ , синяя поверхность) и его касательного пространства  $T_{\mathbf{G}}\mathcal{M}$  (также  $\mathcal{S}$ ) в точке  $\mathbf{G}$ . Логарифмическое отображение (Log) переносит точки с многообразия (например,  $C_1$ ) в векторы касательного пространства (например,  $\xi_1$ ). Экспоненциальное отображение (Exp) выполняет обратную операцию. Зеленые линии  $\delta(C_1, G)$  и  $\delta(C_2, G)$  — геодезические пути на многообразии.

### 4.3 Римановы генеративные модели (RSGM)

Стандартные генеративные модели на основе оценки градиента (SGM), как было показано ранее, по своей конструкции ограничены евклидовыми пространствами. Для преодоления этого ограничения и работы с данными на многообразиях были разработаны римановы генеративные модели (RSGM) [1]. Идея RSGM заключается в том, чтобы определить прямой и обратный диффузионные процессы непосредственно на римановом многообразии  $\mathcal{M}$ , используя его внутреннюю геометрию, описанную в разделе 4.2.

**Прямой процесс на многообразии.** Аналогично евклидову случаю, прямой процесс в RSGM постепенно разрушает структуру данных, преобразуя исходное распределение  $p_0(\mathbf{C})$  на многообразии в простое априорное распределение (например,

в равномерное на компактных многообразиях). Этот процесс описывается стохастическим дифференциальным уравнением (SDE), определенным непосредственно на многообразии  $\mathcal{M}$ :

$$d\mathbf{C}_t = b(\mathbf{C}_t)dt + \sigma(t)d\mathbf{B}_t^{\mathcal{M}}, \quad (12)$$

где  $b(\cdot)$  — векторное поле дрейфа на многообразии,  $\sigma(t)$  — скалярная функция диффузии, а  $d\mathbf{B}_t^{\mathcal{M}}$  — приращение риманова броуновского движения. Риманово броуновское движение представляет собой случайное блуждание, траектория которого ограничена поверхностью многообразия, что схематично показано на рисунке 4.

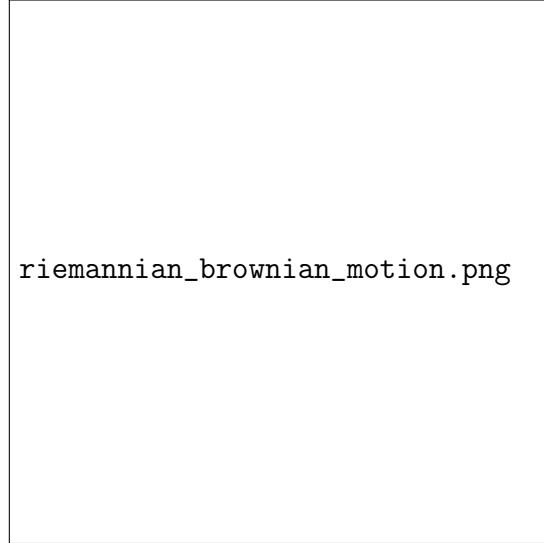


Рис. 3: Концептуальная иллюстрация траектории риманова броуновского движения на многообразии (сфере). В отличие от стандартного броуновского движения, случайное блуждание происходит по поверхности многообразия, следуя его кривизне.

**Обратный процесс на многообразии.** Как и в евклидовом случае, существует обратное по времени SDE, которое позволяет восстанавливать данные из шума. Уравнение для обратного процесса на многообразии  $\mathcal{M}$  имеет следующий вид [1, Theorem 3.1]:

$$d\mathbf{C}_t = \left[ -b(\mathbf{C}_t) + \sigma(t)^2 \nabla_{\mathcal{M}} \log p_t(\mathbf{C}_t) \right] dt + \sigma(t)d\bar{\mathbf{B}}_t^{\mathcal{M}}, \quad (13)$$

где  $\nabla_{\mathcal{M}} \log p_t(\mathbf{C}_t)$  — это *риманова функция оценки* (Riemannian score function), то есть градиент логарифма плотности, вычисленный на многообразии. Это ключевое отличие от стандартных SGM, где используется евклидов градиент.

**Аппроксимация римановой функции оценки.** Основная задача в RSGM, как и в SGM, заключается в аппроксимации неизвестной функции оценки. Для этого используется нейронная сеть  $s_{\theta}(\mathbf{C}, t)$ , которая обучается предсказывать риманов градиент. Поскольку операции на многообразии нелинейны, нейронная сеть, как правило, работает в касательных пространствах. Это означает, что модель  $s_{\theta}$  принимает на вход точку на многообразии  $\mathbf{C}$  и предсказывает вектор в касательном пространстве  $T_{\mathbf{C}}\mathcal{M}$ , который аппроксимирует истинную функцию оценки.

Такой подход гарантирует, что весь процесс генерации происходит на многообразии, а сгенерированные объекты (в нашем случае ковариационные матрицы) по

построению сохраняют свои структурные свойства, такие как симметричность и положительная определенность.

В следующей главе на основе этих концепций будет подробно описан предложенный метод прогнозирования наборов временных рядов, реализующий условную генерацию ковариационных матриц с помощью RSGM.

#### 4.4 Римановы генеративные модели (RSGM)

Стандартные генеративные модели на основе оценки градиента (SGM), как было показано ранее, по своей конструкции ограничены евклидовыми пространствами. Для преодоления этого ограничения и работы с данными на многообразиях были разработаны римановы генеративные модели (RSGM) [1]. Идея RSGM заключается в том, чтобы определить прямой и обратный диффузионные процессы непосредственно на римановом многообразии  $\mathcal{M}$ , используя его внутреннюю геометрию, описанную в разделе 4.2.

**Прямой процесс на многообразии.** Аналогично евклидову случаю, прямой процесс в RSGM постепенно разрушает преобразует исходное распределение  $p_0(\mathbf{C})$  на многообразии в априорное распределение. Этот процесс описывается стохастическим дифференциальным уравнением (SDE), определенным непосредственно на многообразии  $\mathcal{M}$ :

$$d\mathbf{C}_t = b(\mathbf{C}_t)dt + \sigma(t)d\mathbf{B}_t^{\mathcal{M}}, \quad (14)$$

где  $b(\cdot)$  — векторное поле дрейфа на многообразии,  $\sigma(t)$  — скалярная функция диффузии, а  $d\mathbf{B}_t^{\mathcal{M}}$  — приращение риманова броуновского движения. Риманово броуновское движение представляет собой случайное блуждание, траектория которого ограничена поверхностью многообразия, что схематично показано на рисунке 4.

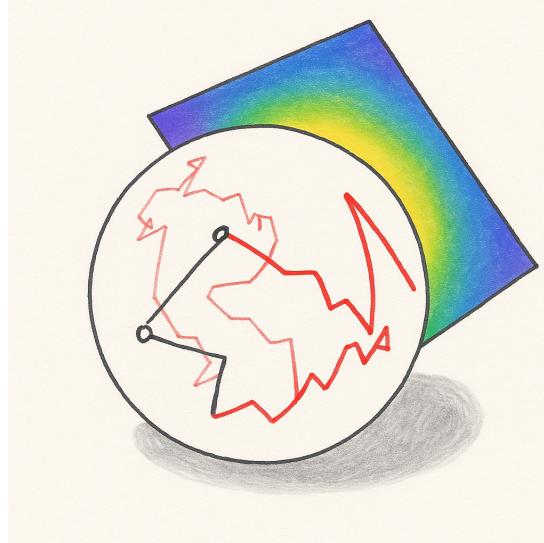


Рис. 4: Концептуальная иллюстрация траектории риманова броуновского движения на многообразии (сфере). В отличие от стандартного броуновского движения, случайное блуждание происходит по поверхности многообразия, следуя его кривизне.

**Обратный процесс на многообразии.** Как и в евклидовом случае, существует обратное по времени SDE, которое позволяет восстанавливать данные из шума.

Уравнение для обратного процесса на многообразии  $\mathcal{M}$  имеет следующий вид [1, Theorem 3.1]:

$$d\mathbf{C}_t = [-b(\mathbf{C}_t) + \sigma(t)^2 \nabla_{\mathcal{M}} \log p_t(\mathbf{C}_t)] dt + \sigma(t) d\bar{\mathbf{B}}_t^{\mathcal{M}}, \quad (15)$$

где  $\nabla_{\mathcal{M}} \log p_t(\mathbf{C}_t)$  — это *риманова функция оценки* (Riemannian score function), то есть градиент логарифма плотности, вычисленный на многообразии. Это ключевое отличие от стандартных SGM, где используется евклидов градиент.

**Аппроксимация римановой функции оценки.** Основная задача в RSGM, как и в SGM, заключается в аппроксимации неизвестной функции оценки. Для этого используется нейронная сеть  $s_{\theta}(\mathbf{C}, t)$ , которая обучается предсказывать риманов градиент. Поскольку операции на многообразии нелинейны, нейронная сеть, как правило, работает в касательных пространствах. Это означает, что модель  $s_{\theta}$  принимает на вход точку на многообразии  $\mathbf{C}$  и предсказывает вектор в касательном пространстве  $T_{\mathbf{C}}\mathcal{M}$ , который аппроксимирует истинную функцию оценки.

Такой подход гарантирует, что весь процесс генерации происходит на многообразии, а сгенерированные объекты (в нашем случае ковариационные матрицы) по построению сохраняют свои структурные свойства, такие как симметричность и положительная определенность.

В следующей главе на основе этих концепций будет подробно описан предложенный метод прогнозирования наборов временных рядов, реализующий условную генерацию ковариационных матриц с помощью RSGM.

## 5 Предложенный метод

В данной главе подробно описывается предложенный метод прогнозирования наборов временных рядов. Общая схема метода состоит из трех этапов: кодирование исходных данных в последовательность ковариационных матриц, прогнозирование в пространстве этих матриц с помощью условной римановой генеративной модели и декодирование результата. Основной вклад и новизна метода заключаются во втором этапе — реализации прогнозирования как задачи условной генерации на многообразии  $Sym^+(n)$ .

### 5.1 Условная генерация ковариационных матриц с помощью RSGM

Задача прогнозирования по своей сути является задачей условной генерации. Необходимо смоделировать условное распределение вероятностей  $p(\mathbf{C}_{t+\Delta} | \mathbf{C}_t)$ , где  $\mathbf{C}_{t+\Delta}$  — будущая (целевая) матрица, а  $\mathbf{C}_t$  — текущая матрица, выступающая в качестве условия.

Для решения этой задачи мы адаптируем фреймворк RSGM. Прямой диффузионный процесс, как и ранее, применяется к целевым данным, то есть к матрице  $\mathbf{C}_{t+\Delta}$ . Однако обратный процесс генерации должен быть обусловлен матрицей  $\mathbf{C}_t$ . Это достигается путем модификации нейронной сети, аппроксимирующей функцию оценки. Вместо  $s_{\theta}(\mathbf{C}, \tau)$ , где  $\tau$  — время диффузии, используется условная модель  $s_{\theta}(\mathbf{C}_{\tau}, \tau, \mathbf{C}_t)$ , которая принимает на вход не только зашумленную матрицу  $\mathbf{C}_{\tau}$  и время  $\tau$ , но и матрицу-условие  $\mathbf{C}_t$ .

Таким образом, обратное SDE для условной генерации (15) принимает вид:

$$d\mathbf{C}_\tau = [-b(\mathbf{C}_\tau, \tau) + g(\tau)^2 s_\theta(\mathbf{C}_\tau, \tau, \mathbf{C}_t)] d\tau + g(\tau) d\bar{\mathbf{B}}_\tau^{\mathcal{M}}. \quad (16)$$

Здесь время  $\tau$  течет от  $T$  до 0. Процесс генерации начинается с сэмпла  $\mathbf{C}_T$  из априорного распределения, и на каждом шаге решения этого уравнения модель использует информацию о  $\mathbf{C}_t$  для корректировки траектории, направляя ее к правдоподобному значению  $\mathbf{C}_{t+\Delta}$ .

## 5.2 Процесс обучения и прогнозирования

Для реализации предложенного метода необходимо определить алгоритмы обучения условной score-сети и последующего использования ее для прогнозирования.

**Обучение модели.** Нейронная сеть  $s_\theta$  обучается на парах матриц  $(\mathbf{C}_t, \mathbf{C}_{t+\Delta})$  из обучающей выборки. Функция потерь является условной версией Denoising Score Matching (7):

$$\mathcal{L}(\theta) = \mathbb{E}_{\tau, (\mathbf{C}_t, \mathbf{C}_{t+\Delta}), \mathbf{C}_{t+\Delta}(\tau)} \left[ \lambda(\tau) \|s_\theta(\mathbf{C}_{t+\Delta}(\tau), \tau, \mathbf{C}_t) - \nabla_{\mathbf{C}_{t+\Delta}(\tau)} \log p_{\tau|0}(\mathbf{C}_{t+\Delta}(\tau) | \mathbf{C}_{t+\Delta})\|_{\mathbf{C}_{t+\Delta}(\tau)}^2 \right], \quad (17)$$

где математическое ожидание берется по парам из набора данных, по времени диффузии  $\tau$  и по запшумленным версиям  $\mathbf{C}_{t+\Delta}(\tau)$ , а норма  $\|\cdot\|_{\mathbf{C}}^2$  является римановой нормой в касательном пространстве. Алгоритм обучения представлен в виде псевдокода в Алгоритме 1.

---

### Algorithm 1 Обучение условной RSGM

---

- 1: **repeat**
  - 2:    Взять случайную пару  $(\mathbf{C}_t, \mathbf{C}_{t+\Delta})$  из обучающей выборки.
  - 3:    Выбрать случайное время диффузии  $\tau \sim \mathcal{U}(0, T)$ .
  - 4:    Применить прямой процесс запшумления к  $\mathbf{C}_{t+\Delta}$  для получения  $\mathbf{C}_{t+\Delta}(\tau)$ .
  - 5:    Вычислить истинный score (градиент)  $\nabla_{\mathbf{C}_{t+\Delta}(\tau)} \log p_{\tau|0}(\mathbf{C}_{t+\Delta}(\tau) | \mathbf{C}_{t+\Delta})$ .
  - 6:    Вычислить предсказание модели  $s_\theta(\mathbf{C}_{t+\Delta}(\tau), \tau, \mathbf{C}_t)$ .
  - 7:    Вычислить ошибку согласно (17).
  - 8:    Выполнить шаг градиентного спуска для обновления параметров  $\theta$ .
  - 9: **until** условие остановки выполнено
- 

**Прогнозирование (семплирование).** После обучения модели, процесс прогнозирования для новой матрицы  $\mathbf{C}_t$  выполняется путем решения обратного SDE (16). Процедура, представленная в Алгоритме 2, позволяет сгенерировать одну реализацию прогноза  $\hat{\mathbf{C}}_{t+\Delta}$ .

Описанный подход позволяет моделировать сложные условные зависимости в пространстве ковариационных матриц, что является ключевым для точного прогнозирования наборов временных рядов.

---

**Algorithm 2** Прогнозирование с помощью условной RSGM

---

**Require:** Обученная модель  $s_\theta$ , матрица-условие  $\mathbf{C}_t$ , число шагов  $N$ .

```
1: Взять начальный семпл из априорного распределения:  $\hat{\mathbf{C}}_N \sim p_T$ .
2: for  $i = N, \dots, 1$  do
3:   Вычислить шаг обратного SDE для получения  $\hat{\mathbf{C}}_{i-1}$  из  $\hat{\mathbf{C}}_i$ , используя
    $s_\theta(\hat{\mathbf{C}}_i, \tau_i, \mathbf{C}_t)$ .
4: end for
5: return  $\hat{\mathbf{C}}_{t+\Delta} = \hat{\mathbf{C}}_0$ .
```

---

## 6 Вычислительный эксперимент

### 6.1 Экспериментальные данные

## 7 Заключение

## Список литературы

- [1] Valentin De Bortoli, Émile Mathieu, Michael Hutchinson, James Thornton, Yee Whye Teh, and Arnaud Doucet. Riemannian score-based generative modelling. In *Advances in Neural Information Processing Systems*, 2022.
- [2] James B Elsner and Anastasios A Tsonis. *Singular Spectrum Analysis: A New Tool in Time Series Analysis*. Plenum Press, 1996.
- [3] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.
- [4] James Durbin and Siem Jan Koopman. *Time Series Analysis by State Space Methods*. Oxford University Press, 2012.
- [5] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, 2020.
- [6] Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. In *Advances in Neural Information Processing Systems*, 2019.
- [7] Prafulla Dhariwal and Alex Nichol. Diffusion models beat gans on image synthesis. In *Advances in Neural Information Processing Systems*, 2021.
- [8] Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations (ICLR)*, 2021.
- [9] Brian D. O. Anderson. Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, 12(3):313–326, 1982.
- [10] Pascal Vincent. A connection between score matching and denoising autoencoders. *Neural Computation*, 23(7):1661–1674, 2011.

- [11] Xavier Pennec. Intrinsic statistics on riemannian manifolds: Basic tools for geometric measurements. volume 25, pages 127–154, 2006.
- [12] P. Thomas Fletcher and Sarang Joshi. Riemannian geometry for the analysis of spherical images. *IEEE Transactions on Medical Imaging*, 26(7):955–962, 2007.