

Учебный Модуль: Линейная Алгебра для Обучения с Подкреплением

Часть 1: Теоретические Основы

Это первая часть учебного модуля, посвященная теоретическим основам линейной алгебры, которые являются ключевыми для понимания и разработки алгоритмов обучения с подкреплением. Материал изложен с расчетом на студентов первого курса технических ВУЗов и стремится обеспечить как фундаментальное понимание, так и связь с практическими аспектами RL.

расстояний между векторами состояний, для регуляризации в моделях машинного обучения (например, L1 и L2 регуляризация весов нейронной сети), или для определения условий сходимости алгоритмов.

1.4. Линейная Комбинация Векторов

Линейной комбинацией векторов $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ с коэффициентами c_1, c_2, \dots, c_m (где $c_i \in \mathbb{R}$) называется вектор \vec{w} , равный:

$$\vec{w} = c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_m\vec{v}_m = \sum_{i=1}^m c_i\vec{v}_i$$

Например, если $\vec{v}_1 = (1, 0)$ и $\vec{v}_2 = (0, 1)$, то любая линейная комбинация $c_1\vec{v}_1 + c_2\vec{v}_2 = (c_1, c_2)$ может представить любой вектор на плоскости \mathbb{R}^2 .

В RL, представление функции ценности $V(s)$ или Q-функции $Q(s, a)$ часто аппроксимируется как линейная комбинация базисных функций (или признаков) $\phi_i(s)$:

$$V(s) \approx \sum_{i=1}^k w_i \phi_i(s) = \vec{w}^T \vec{\phi}(s)$$

Здесь $\vec{w} = (w_1, \dots, w_k)$ – вектор весов, а $\vec{\phi}(s) = (\phi_1(s), \dots, \phi_k(s))$ – вектор признаков состояния s . Обучение такой модели заключается в нахождении оптимальных весов w_i .

1.5. Линейная Зависимость и Независимость Векторов

Система векторов $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ называется **линейно зависимой**, если существуют такие скаляры c_1, c_2, \dots, c_m , не все равные нулю одновременно, что их линейная комбинация равна нулевому вектору:

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_m\vec{v}_m = \vec{0}, \quad \text{где хотя бы один } c_i \neq 0$$

Если это равенство выполняется только тогда, когда все коэффициенты $c_1 = c_2 = \dots = c_m = 0$, то система векторов называется **линейно независимой**.

Интерпретация:

- **Линейно зависимые векторы:** По крайней мере один из векторов системы может быть выражен как линейная комбинация остальных. Например, если $c_1 \neq 0$, то $\vec{v}_1 = -\frac{c_2}{c_1}\vec{v}_2 - \dots - \frac{c_m}{c_1}\vec{v}_m$. Геометрически, для двух векторов на плоскости, линейная зависимость означает, что они коллинеарны (лежат на одной прямой). Для трех векторов в пространстве – что они компланарны (лежат в одной плоскости).
- **Линейно независимые векторы:** Ни один из векторов системы не может быть выражен как линейная комбинация остальных. Они

указывают на "разные" направления, и ни один из них не является избыточным.

Проверка на линейную зависимость/независимость:

Для системы векторов, заданных своими координатами, вопрос о линейной зависимости сводится к решению системы линейных однородных уравнений. Если эта система имеет ненулевое решение, векторы линейно зависимы. Если единственное решение – нулевое, то векторы линейно независимы. Также можно использовать определитель матрицы, составленной из координат векторов (если число векторов равно размерности пространства).

В RL понимание линейной независимости важно, например, при выборе признаков для представления состояний. Если признаки линейно зависимы, это может привести к избыточности информации и проблемам с обучением моделей (например, мультиколлинеарность в линейных моделях).

1.6. Базис и Размерность Векторного Пространства

Базисом векторного пространства V называется упорядоченный набор линейно независимых векторов $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ такой, что любой вектор $\vec{v} \in V$ может быть единственным образом представлен в виде их линейной комбинации:

$$\vec{v} = x_1\vec{e}_1 + x_2\vec{e}_2 + \dots + x_n\vec{e}_n$$

Числа (x_1, x_2, \dots, x_n) называются **координатами вектора** \vec{v} в базисе $\{\vec{e}_i\}$.

Размерностью векторного пространства V , обозначаемой $\dim V$, называется количество векторов в любом его базисе. Если пространство не имеет конечного базиса, оно называется бесконечномерным.

Стандартный (или канонический) базис в \mathbb{R}^n состоит из векторов:

$$\vec{e}_1 = (1, 0, \dots, 0)$$

$$\vec{e}_2 = (0, 1, \dots, 0)$$

...

$$\vec{e}_n = (0, 0, \dots, 1)$$

В этом базисе координаты любого вектора $\vec{a} = (a_1, a_2, \dots, a_n)$ совпадают с его компонентами.

Пример:

На плоскости (\mathbb{R}^2) любые два неколлинеарных вектора образуют базис. Размерность \mathbb{R}^2 равна 2.

В трехмерном пространстве (\mathbb{R}^3) любые три некомпланарных вектора образуют базис. Размерность \mathbb{R}^3 равна 3.

Выбор подходящего базиса может значительно упростить решение задач в RL.

Например, методы понижения размерности, такие как анализ главных компонент (PCA), ищут новый базис, в котором данные наилучшим образом представлены меньшим числом координат, сохраняя при этом максимум информации. Это помогает бороться с "проклятием размерности" в RL.

Связь с обучением с подкреплением:

- **Представление состояний:** Векторы состояний s существуют в некотором пространстве состояний S . Если S можно представить как векторное пространство, то выбор базиса для этого пространства может быть важен для эффективности алгоритмов RL.
- **Аппроксимация функций:** При аппроксимации функций ценности или стратегий с помощью линейных моделей (например, $V(s) \approx \vec{w}^T \vec{\phi}(s)$), набор базисных функций $\{\phi_i(s)\}$ можно рассматривать как формирующий базис в пространстве функций. Линейная независимость этих базисных функций важна для стабильности и единственности решения.
- **Политики (Policies):** В некоторых случаях политика $\pi(a|s)$ может быть представлена как вектор в пространстве распределений вероятностей, и операции линейной алгебры могут применяться для их анализа или обновления.

Понимание этих концепций векторов является первым шагом к освоению более сложных тем линейной алгебры, таких как матрицы, определители, собственные значения и векторы, которые также играют ключевую роль в теории и практике обучения с подкреплением.

Раздел 2: Матрицы в Линейной Алгебре и их Роль в Обучении с Подкреплением

2.1. Определение и Типы Матриц

Матрицей размера $n \times m$ (читается "n на m") называется прямоугольная таблица чисел, состоящая из n строк и m столбцов. Числа, составляющие матрицу, называются ее **элементами**.

Элемент матрицы, стоящий на пересечении i -й строки и j -го столбца, обозначается как a_{ij} , где i – номер строки, а j – номер столбца. Сами матрицы обычно обозначаются заглавными латинскими буквами, например, A, B, C .

Общий вид матрицы A размера $n \times m$:

$$A = A_{n \times m} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}$$

Основные типы матриц:

1. **Вектор-строка (матрица-строка):** Матрица, состоящая из одной строки ($n = 1$).

Пример: $A = (2 \quad -1 \quad 3 \quad 5)$

2. **Вектор-столбец (матрица-столбец):** Матрица, состоящая из одного столбца ($m = 1$).

Пример: $B = \begin{pmatrix} 1 \\ 4 \\ -8 \end{pmatrix}$

3. **Квадратная матрица:** Матрица, у которой число строк равно числу столбцов ($n = m$). Число n называется **порядком** квадратной матрицы.

Элементы $a_{11}, a_{22}, \dots, a_{nn}$ образуют **главную диагональ** квадратной матрицы.

Элементы $a_{1n}, a_{2,n-1}, \dots, a_{n1}$ образуют **побочную диагональ**.

4. **Нулевая матрица (O):** Матрица, все элементы которой равны нулю. Обозначается $O_{n \times m}$ или просто O .

Пример: $O_{2 \times 3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

5. **Диагональная матрица:** Квадратная матрица, у которой все элементы вне главной диагонали равны нулю.

Пример: $D = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$

6. **Единичная матрица (I или E):** Диагональная матрица, у которой все элементы на главной диагонали равны единице. Обозначается I_n или E_n (или просто I, E , если порядок ясен из контекста).

Пример: $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Единичная матрица играет роль единицы в умножении матриц: $AI = IA = A$.

7. **Треугольная матрица:**

- **Верхняя треугольная матрица:** Квадратная матрица, у которой все элементы ниже главной диагонали равны нулю ($a_{ij} = 0$ при $i > j$).

- **Нижняя треугольная матрица:** Квадратная матрица, у которой все элементы выше главной диагонали равны нулю ($a_{ij} = 0$ при $i < j$).

8. **Симметрическая матрица:** Квадратная матрица, у которой элементы, симметричные относительно главной диагонали, равны ($a_{ij} = a_{ji}$ для всех i, j). Это означает, что $A = A^T$ (где A^T – транспонированная матрица, см. ниже).

Пример: $S = \begin{pmatrix} 2 & 1 & 5 \\ 1 & 3 & 6 \\ 5 & 6 & 4 \end{pmatrix}$

9. **Кососимметрическая (антисимметрическая) матрица:** Квадратная матрица, у которой элементы, симметричные относительно главной диагонали, противоположны по знаку ($a_{ij} = -a_{ji}$), а элементы на главной диагонали равны нулю ($a_{ii} = 0$). Это означает, что $A = -A^T$.

Равенство матриц: Две матрицы A и B называются равными ($A = B$), если они имеют одинаковые размеры ($n \times m$) и их соответствующие элементы равны ($a_{ij} = b_{ij}$ для всех i, j).

Применение матриц в RL:

- **Матрицы переходных вероятностей (Transition Probability Matrix):** В Марковских процессах принятия решений (MDP) матрица P определяет вероятности перехода из одного состояния s в другое состояние s' при совершении действия a : $P_{ss'}^a = P(s_{t+1} = s' | s_t = s, a_t = a)$. Это фундаментальное понятие для многих алгоритмов RL.
- **Матрицы вознаграждений (Reward Matrix):** Иногда вознаграждения также представляются в виде матриц или векторов, связанных с состояниями и/или действиями.
- **Представление данных:** Наборы данных, такие как батчи траекторий (состояние, действие, вознаграждение, следующее состояние), могут быть организованы в виде матриц для обработки.
- **Весовые матрицы в нейронных сетях:** В глубоком обучении с подкреплением (Deep RL) нейронные сети используются для аппроксимации функций ценности или стратегий. Веса между слоями нейронов хранятся в матрицах.
- **Системы линейных уравнений:** Многие задачи в RL, особенно при табличных методах или линейной аппроксимации, сводятся к решению систем линейных уравнений, которые удобно представлять и решать с использованием матриц (например, уравнения Беллмана).

2.2. Операции над Матрицами

2.2.1. Сложение и Вычитание Матриц

Сложение (и вычитание) матриц определено только для матриц **одинакового размера**.

Суммой двух матриц $A = (a_{ij})$ и $B = (b_{ij})$ размера $n \times m$ называется матрица $C = (c_{ij})$ того же размера $n \times m$, элементы которой равны суммам соответствующих элементов матриц A и B :

$$c_{ij} = a_{ij} + b_{ij}$$

Записывают: $C = A + B$.

Разностью двух матриц A и B одинакового размера называется матрица $D = (d_{ij})$ того же размера, элементы которой равны разностям соответствующих элементов матриц A и B :

$$d_{ij} = a_{ij} - b_{ij}$$

Записывают: $D = A - B$. Это эквивалентно $A + (-B)$, где $-B$ - матрица, полученная умножением B на -1 .

Свойства сложения матриц:

1. Коммутативность: $A + B = B + A$
2. Ассоциативность: $(A + B) + C = A + (B + C)$
3. Существование нулевой матрицы: $A + O = A$ (где O – нулевая матрица того же размера, что и A)
4. Существование противоположной матрицы: для каждой матрицы A существует матрица $-A$ (элементы которой противоположны элементам A) такая, что $A + (-A) = O$.

2.2.2. Умножение Матрицы на Скаляр (Число)

Произведением матрицы $A = (a_{ij})$ размера $n \times m$ на скаляр (число) $k \in \mathbb{R}$ называется матрица $C = (c_{ij})$ того же размера $n \times m$, каждый элемент которой равен произведению соответствующего элемента матрицы A на скаляр k :

$$c_{ij} = k \cdot a_{ij}$$

Записывают: $C = kA$.

Свойства умножения матрицы на скаляр:

1. Ассоциативность относительно скалярных множителей: $(k_1 k_2)A = k_1(k_2 A)$
2. Дистрибутивность относительно суммы скаляров: $(k_1 + k_2)A = k_1 A + k_2 A$
3. Дистрибутивность относительно суммы матриц: $k(A + B) = kA + kB$
4. Умножение на единичный скаляр: $1 \cdot A = A$
5. Умножение на нулевой скаляр: $0 \cdot A = O$

2.2.3. Произведение Матриц

Умножение двух матриц A и B определено только в том случае, если **число столбцов первой матрицы (A) равно числу строк второй матрицы (B)**. Такие матрицы называются **согласованными** для умножения.

Пусть матрица A имеет размер $n \times s$, а матрица B имеет размер $s \times m$. Тогда их **произведением** $C = AB$ будет матрица размера $n \times m$, элемент c_{ij} которой (стоящий в i -й строке и j -м столбце) равен сумме произведений элементов i -й строки матрицы A на соответствующие элементы j -го столбца матрицы B :

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{is}b_{sj} = \sum_{k=1}^s a_{ik}b_{kj}$$

Правило умножения "строка на столбец": Чтобы получить элемент c_{ij} матрицы-произведения, нужно взять i -ю строку первой матрицы и j -й столбец второй матрицы, перемножить их соответствующие элементы и сложить полученные произведения.

Пример:

Пусть $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ (размер 2×2) и $B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$ (размер 2×2).

Тогда $C = AB = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$, где:

$$c_{11} = 1 \cdot 5 + 2 \cdot 7 = 5 + 14 = 19$$

$$c_{12} = 1 \cdot 6 + 2 \cdot 8 = 6 + 16 = 22$$

$$c_{21} = 3 \cdot 5 + 4 \cdot 7 = 15 + 28 = 43$$

$$c_{22} = 3 \cdot 6 + 4 \cdot 8 = 18 + 32 = 50$$

$$\text{Итак, } C = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}.$$

Свойства умножения матриц:

- Некоммутативность (в общем случае):** $AB \neq BA$. Даже если оба произведения AB и BA определены (например, для квадратных матриц одного порядка), они не всегда равны. Если $AB = BA$, то матрицы A и B называются **перестановочными** или **коммутирующими**.
- Ассоциативность:** $(AB)C = A(BC)$ (если размеры матриц позволяют выполнить операции).
- Дистрибутивность:**
 - $A(B + C) = AB + AC$ (левая дистрибутивность)
 - $(A + B)C = AC + BC$ (правая дистрибутивность)
- Умножение на единичную матрицу:** Если A – матрица $n \times m$, то $AI_m = A$ и $I_n A = A$, где I_m и I_n – единичные матрицы соответствующих порядков.
- Умножение на нулевую матрицу:** $AO = O$ и $OA = O$ (где O – нулевые матрицы подходящих размеров).
- $(kA)B = A(kB) = k(AB)$ для любого скаляра k .

Применение умножения матриц в RL:

- **Обновление вектора ценности состояния (Value Iteration):** В алгоритме итерации по ценностям, обновление вектора ценностей $V_{k+1}(s)$ для всех состояний s можно записать в матричной форме. Например, для детерминированной политики π :
 $V_{k+1} = R^\pi + \gamma P^\pi V_k$, где V_k – вектор ценностей на k -й итерации, R^π – вектор ожидаемых вознаграждений, P^π – матрица переходных вероятностей для политики π , γ – фактор дисконтирования. Здесь происходит умножение матрицы на вектор (вектор рассматривается как матрица-столбец).
- **Последовательные преобразования в нейронных сетях:** Прямое распространение сигнала в многослойной нейронной сети включает последовательное умножение вектора активаций предыдущего слоя на матрицу весов текущего слоя и применение функции активации.
- **Представление линейных преобразований:** Матрицы используются для представления линейных преобразований векторов. Например, если состояние s представлено вектором, то линейное преобразование этого состояния $s' = As$ выполняется умножением на матрицу A .

2.2.4. Транспонирование Матрицы

Транспонированной матрицей A^T (или A') для матрицы $A = (a_{ij})$ размера $n \times m$ называется матрица размера $m \times n$, полученная из A заменой каждой ее строки соответствующим столбцом (или, что то же самое, каждого столбца – соответствующей строкой).

Если $A = (a_{ij})$, то $A^T = (a_{ji}^T)$, где $a_{ji}^T = a_{ij}$.

Пример:

Если $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$, то $A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$.

Свойства транспонирования:

1. $(A^T)^T = A$
2. $(A + B)^T = A^T + B^T$
3. $(kA)^T = kA^T$
4. $(AB)^T = B^T A^T$ (обратите внимание на изменение порядка сомножителей!)

Применение транспонирования в RL:

- **Скалярное произведение векторов:** Скалярное произведение двух векторов-столбцов \vec{x} и \vec{y} можно записать как $\vec{x}^T \vec{y}$.
- **Квадратичные формы:** Выражения вида $\vec{x}^T A \vec{x}$ (квадратичные формы) часто встречаются в оптимизационных задачах и анализе устойчивости.
- **Вычисления градиентов в нейронных сетях:** При обратном распространении ошибки (backpropagation) используются транспонированные матрицы весов.

- **Нормальные уравнения в линейной регрессии:** При решении задач линейной регрессии методом наименьших квадратов (например, для аппроксимации функции ценности) возникают нормальные уравнения вида $X^T X \vec{w} = X^T \vec{y}$, где X – матрица признаков.

2.3. Определитель (Детерминант) Матрицы

Определитель (или детерминант) – это скалярная величина, которая может быть вычислена для каждой **квадратной матрицы** и характеризует многие ее свойства (например, обратимость). Обозначается как $\det(A)$, $|A|$, или $\Delta(A)$.

2.3.1. Определитель матрицы 1-го порядка

Для матрицы $A = (a_{11})$ (один элемент), определитель равен этому элементу:

$$\det(A) = a_{11}$$

2.3.2. Определитель матрицы 2-го порядка

Для квадратной матрицы $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ второго порядка, определитель вычисляется по формуле:

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Это разность произведений элементов главной диагонали и побочной диагонали.

2.3.3. Определитель матрицы 3-го порядка

Для квадратной матрицы $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ третьего порядка, определитель можно вычислить по **правилу Саррюса (правилу треугольников)**:

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

Это правило удобно запомнить так: три слагаемых со знаком "плюс" – это произведения элементов главной диагонали и элементов, образующих треугольники с основаниями, параллельными главной диагонали. Три слагаемых со знаком "минус" – это произведения элементов побочной диагонали и элементов, образующих треугольники с основаниями, параллельными побочной диагонали.

2.3.4. Миноры и Алгебраические Дополнения

Для вычисления определителей матриц порядка $n > 3$ (и как общий метод для $n \geq 2$) используется разложение по строке или столбцу.

- **Минором** M_{ij} элемента a_{ij} квадратной матрицы A называется определитель матрицы $(n - 1)$ -го порядка, полученной из A вычеркиванием i -й строки и j -го столбца.
- **Алгебраическим дополнением** A_{ij} элемента a_{ij} называется его минор, взятый со знаком $(-1)^{i+j}$:

$$A_{ij} = (-1)^{i+j} M_{ij}$$

2.3.5. Разложение Определителя по Строке или Столбцу (Теорема Лапласа)

Определитель квадратной матрицы A равен сумме произведений элементов любой ее строки (или столбца) на их алгебраические дополнения.

- **Разложение по i -й строке:**

$$\det(A) = \sum_{j=1}^n a_{ij} A_{ij} = a_{i1} A_{i1} + a_{i2} A_{i2} + \dots + a_{in} A_{in}$$

- **Разложение по j -му столбцу:**

$$\det(A) = \sum_{i=1}^n a_{ij} A_{ij} = a_{1j} A_{1j} + a_{2j} A_{2j} + \dots + a_{nj} A_{nj}$$

Этот метод позволяет свести вычисление определителя n -го порядка к вычислению n определителей $(n - 1)$ -го порядка, и так далее, пока не дойдем до определителей 2-го или 1-го порядка.

2.3.6. Свойства Определителей

1. Определитель матрицы не изменяется при ее транспонировании: $\det(A) = \det(A^T)$. (Это означает, что все свойства, справедливые для строк, справедливы и для столбцов).
2. Если матрица содержит нулевую строку (или столбец), то ее определитель равен нулю.
3. Если одна из строк (столбцов) матрицы умножается на скаляр k , то определитель матрицы умножается на k .

$$\begin{vmatrix} ka_{11} & ka_{12} \\ a_{21} & a_{22} \end{vmatrix} = k \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

Следствие: Общий множитель элементов одной строки (столбца) можно вынести за знак определителя.

4. При перестановке двух строк (столбцов) матрицы ее определитель меняет знак на противоположный.
5. Если матрица имеет две одинаковые или пропорциональные строки (столбца), то ее определитель равен нулю.

6. Если одна из строк (столбцов) матрицы является линейной комбинацией других строк (столбцов), то ее определитель равен нулю. (Это связано с понятием линейной зависимости строк/столбцов).
7. Определитель не изменится, если к элементам одной строки (столбца) прибавить соответствующие элементы другой строки (столбца), умноженные на одно и то же число.
Это свойство часто используется для упрощения вычисления определителей путем получения нулей в какой-либо строке или столбце.
8. Определитель произведения двух квадратных матриц одного порядка равен произведению их определителей: $\det(AB) = \det(A) \cdot \det(B)$.
9. Определитель треугольной (в частности, диагональной) матрицы равен произведению элементов ее главной диагонали.

Особые случаи:

- **Вырожденная (сингулярная) матрица:** Квадратная матрица A называется вырожденной, если ее определитель равен нулю ($\det(A) = 0$). В противном случае (если $\det(A) \neq 0$), матрица называется **невырожденной (несингулярной)**. Вырожденные матрицы не имеют обратной матрицы. Строки (и столбцы) вырожденной матрицы линейно зависимы.

Применение определителей в RL (и связанных областях):

- **Проверка обратимости матриц:** Важно при решении систем линейных уравнений (например, уравнений Беллмана в матричной форме). Если матрица системы невырождена, система имеет единственное решение.
- **Геометрическая интерпретация:** Абсолютная величина определителя матрицы линейного преобразования показывает, во сколько раз изменяется "объем" (площадь для 2D, объем для 3D) при этом преобразовании. Знак определителя связан с изменением ориентации.
- **Нахождение собственных значений:** Определители используются при составлении характеристического уравнения для нахождения собственных значений матрицы ($\det(A - \lambda I) = 0$).
- **Анализ устойчивости систем:** В теории управления (тесно связанной с RL) определители матриц используются для анализа устойчивости динамических систем.
- **Вычисление ранга матрицы:** Хотя ранг определяется через линейную независимость, он также связан с минорами (определителями подматриц).

2.4. Ранг Матрицы

Ранг матрицы A (обозначается $\text{rank}(A)$, $\text{rg}(A)$, или $r(A)$) — это одна из важнейших числовых характеристик матрицы. Существует несколько эквивалентных определений ранга:

1. **Ранг как максимальное число линейно независимых строк:** Ранг матрицы равен максимальному числу линейно независимых строк матрицы.
2. **Ранг как максимальное число линейно независимых столбцов:** Ранг матрицы равен максимальному числу линейно независимых столбцов матрицы. (Важно: максимальное число линейно независимых строк всегда равно максимальному числу линейно независимых столбцов).
3. **Ранг как порядок наибольшего ненулевого минора:** Ранг матрицы равен наибольшему порядку k ее миноров (определителей квадратных подматриц), отличных от нуля. Если все миноры порядка k равны нулю, то ранг матрицы меньше k . Ранг нулевой матрицы равен нулю.

Свойства ранга матрицы:

1. Ранг матрицы $A_{n \times m}$ не превышает меньшего из ее размеров: $\text{rank}(A) \leq \min(n, m)$.
2. $\text{rank}(A) = \text{rank}(A^T)$.
3. Ранг матрицы не изменяется при элементарных преобразованиях строк или столбцов (см. ниже).
4. Ранг диагональной матрицы равен числу ненулевых элементов на ее главной диагонали.
5. Ранг единичной матрицы I_n равен n .
6. Если матрица A квадратная порядка n , то A невырождена ($\det(A) \neq 0$) тогда и только тогда, когда $\text{rank}(A) = n$. Если $\text{rank}(A) < n$, то матрица вырождена ($\det(A) = 0$).
7. Ранг произведения матриц: $\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B))$.
8. Ранг суммы матриц: $\text{rank}(A + B) \leq \text{rank}(A) + \text{rank}(B)$.

Методы нахождения ранга:

1. Метод элементарных преобразований (метод Гаусса):

Элементарные преобразования строк (столбцов) матрицы – это:

- Перестановка двух строк (столбцов).
- Умножение строки (столбца) на ненулевое число.
- Прибавление к одной строке (столбцу) другой строки (столбца), умноженной на число.

С помощью элементарных преобразований матрицу можно привести к **ступенчатому виду** (или трапециевидному виду). Ранг ступенчатой матрицы равен количеству ее ненулевых строк. Поскольку элементарные преобразования не меняют ранг, ранг исходной матрицы равен рангу полученной ступенчатой матрицы.

2. Метод окаймляющих миноров:

Находят какой-либо ненулевой минор 1-го порядка (т.е. ненулевой элемент). Затем рассматривают миноры 2-го порядка, окаймляющие его (т.е. содержащие его). Если все они равны нулю, ранг равен 1. Если есть ненулевой минор 2-го порядка,

переходят к минорам 3-го порядка, окаймляющим его, и т.д. Ранг равен порядку последнего найденного ненулевого окаймляющего минора.

Применение ранга матрицы в RL:

- **Анализ систем линейных уравнений (Теорема Кронекера-Капелли):** Система линейных уравнений $Ax = b$ совместна (т.е. имеет хотя бы одно решение) тогда и только тогда, когда ранг основной матрицы системы A равен рангу расширенной матрицы $[A|b]$. Если система совместна и $\text{rank}(A) = n$ (где n – число неизвестных), то система имеет единственное решение. Если $\text{rank}(A) < n$, то система имеет бесконечно много решений.
Это важно при решении уравнений Беллмана.
- **Линейная зависимость/независимость векторов:** Ранг матрицы, составленной из векторов-строк (или векторов-столбцов), равен максимальному числу линейно независимых векторов в этой системе.
- **Понижение размерности (Dimensionality Reduction):** В RL часто работают с пространствами состояний высокой размерности. Методы понижения размерности, такие как Анализ Главных Компонент (PCA) или Сингулярное Разложение (SVD), тесно связаны с рангом матрицы данных. Низкоранговая аппроксимация матрицы может помочь выявить наиболее важные признаки и уменьшить вычислительную сложность.
- **Проверка на избыточность признаков:** Если матрица признаков имеет ранг меньше числа признаков, это указывает на линейную зависимость между признаками (мультиколлинеарность), что может быть проблемой для некоторых моделей обучения.

2.5. Обратная Матрица

Обратная матрица A^{-1} для квадратной матрицы A порядка n – это такая матрица, которая при умножении на A (как слева, так и справа) дает единичную матрицу I_n :

$$AA^{-1} = A^{-1}A = I_n$$

Условия существования обратной матрицы:

Обратная матрица A^{-1} существует тогда и только тогда, когда матрица A **невырождена** (т.е. ее определитель $\det(A) \neq 0$).

Если $\det(A) = 0$, матрица A называется вырожденной (сингулярной), и для нее обратной матрицы не существует.

Формула для нахождения обратной матрицы:

Если A – невырожденная квадратная матрица, то ее обратная матрица A^{-1} вычисляется по формуле:

$$A^{-1} = \frac{1}{\det(A)} A_{adj}^T = \frac{1}{\det(A)} C^T$$

где:

- $\det(A)$ – определитель матрицы A .
- A_{adj} (или C) – **матрица алгебраических дополнений** (союзная или присоединенная матрица), состоящая из алгебраических дополнений A_{ij} элементов a_{ij} матрицы A . То есть, $C = (A_{ij})$.
- C^T (или A_{adj}^T) – транспонированная матрица алгебраических дополнений (также называется **присоединенной матрицей** или **союзной матрицей**).

Алгоритм нахождения обратной матрицы:

1. Вычислить определитель матрицы A , $\det(A)$. Если $\det(A) = 0$, то обратной матрицы не существует. Прекратить вычисления.
2. Если $\det(A) \neq 0$, найти все алгебраические дополнения A_{ij} для каждого элемента a_{ij} матрицы A .
3. Составить матрицу алгебраических дополнений $C = (A_{ij})$.
4. Транспонировать матрицу C для получения присоединенной матрицы C^T .
5. Умножить каждый элемент матрицы C^T на $\frac{1}{\det(A)}$. Полученная матрица и будет A^{-1} .

Свойства обратной матрицы:

1. Обратная матрица единственна (если существует).
2. $(A^{-1})^{-1} = A$
3. $(AB)^{-1} = B^{-1}A^{-1}$ (если A и B невырождены и одного порядка; обратите внимание на изменение порядка сомножителей).
4. $(A^T)^{-1} = (A^{-1})^T$
5. $(kA)^{-1} = \frac{1}{k}A^{-1}$ (для $k \neq 0$).
6. $\det(A^{-1}) = \frac{1}{\det(A)} = (\det(A))^{-1}$.

Метод Гаусса-Жордана для нахождения обратной матрицы:

Другой распространенный метод – это метод Гаусса-Жордана. К исходной матрице A справа приписывают единичную матрицу I того же порядка, получая расширенную матрицу $[A|I]$. Затем с помощью элементарных преобразований строк приводят левую часть (матрицу A) к единичной матрице I . В результате этих же преобразований на месте правой части (исходной I) окажется обратная матрица A^{-1} . То есть,

$$[A|I] \xrightarrow{\text{эл. преобр.}} [I|A^{-1}].$$

Применение обратной матрицы в RL:

- **Решение систем линейных уравнений:** Если система линейных уравнений представлена в матричной форме $Ax = b$ и матрица A квадратная и невырожденная, то решение можно найти как $x = A^{-1}b$. Это напрямую применимо к

решению уравнений Беллмана для функции ценности в табличных методах RL:
 $V = (I - \gamma P^\pi)^{-1} R^\pi$.

- **Метод наименьших квадратов (Least Squares):** В задачах линейной аппроксимации (например, аппроксимация функции ценности $V(s) \approx \vec{w}^T \vec{\phi}(s)$), вектор весов \vec{w} часто находится решением нормальных уравнений $(X^T X) \vec{w} = X^T \vec{y}$. Если матрица $X^T X$ невырождена, то $\vec{w} = (X^T X)^{-1} X^T \vec{y}$. Выражение $(X^T X)^{-1} X^T$ называется псевдообратной матрицей для X (если X не квадратная).
- **Вычисление ковариационных матриц и их обратных (матриц точности):** В вероятностных моделях и некоторых алгоритмах RL (например, основанных на Гауссовских процессах или при работе с многомерными нормальными распределениями) используются ковариационные матрицы. Обратная ковариационная матрица (матрица точности) также играет важную роль.
- **Изменение базиса:** Обратные матрицы используются при переходе от одного базиса к другому.

Понимание матриц, их свойств и операций над ними является критически важным для освоения большинства алгоритмов и теоретических концепций в обучении с подкреплением.

Раздел 3: Скалярное Произведение Векторов и его Значение в Обучении с Подкреплением

3.1. Определение Скалярного Произведения

Скалярным произведением (иногда называемым внутренним произведением) двух векторов \vec{a} и \vec{b} в евклидовом пространстве называется **скалярная величина** (число), которая может быть определена несколькими эквивалентными способами.

3.1.1. Геометрическое определение

Скалярное произведение векторов \vec{a} и \vec{b} равно произведению их длин (модулей) на косинус угла θ между ними:

$$\vec{a} \cdot \vec{b} = \|\vec{a}\| \cdot \|\vec{b}\| \cdot \cos(\theta)$$

где $\|\vec{a}\|$ и $\|\vec{b}\|$ – длины векторов \vec{a} и \vec{b} соответственно, а θ – угол между векторами ($0 \leq \theta \leq \pi$).

Геометрический смысл:

- Скалярное произведение $\|\vec{a}\|(\|\vec{b}\| \cos \theta)$ можно интерпретировать как произведение длины вектора \vec{a} на проекцию вектора \vec{b} на направление вектора \vec{a} (или наоборот: $\|\vec{b}\|(\|\vec{a}\| \cos \theta)$ – произведение длины \vec{b} на проекцию \vec{a} на \vec{b}).
- Если угол $\theta < 90^\circ$ (острый), то $\cos \theta > 0$, и скалярное произведение положительно.

- Если угол $\theta = 90^\circ$ (прямой), то $\cos \theta = 0$, и скалярное произведение равно нулю. В этом случае векторы называются **ортогональными** (перпендикулярными).
- Если угол $\theta > 90^\circ$ (тупой), то $\cos \theta < 0$, и скалярное произведение отрицательно.
- Если векторы коллинеарны и сонаправлены ($\theta = 0^\circ$), то $\cos \theta = 1$, и $\vec{a} \cdot \vec{b} = \|\vec{a}\| \cdot \|\vec{b}\|$.
- Если векторы коллинеарны и противоположно направлены ($\theta = 180^\circ$), то $\cos \theta = -1$, и $\vec{a} \cdot \vec{b} = -\|\vec{a}\| \cdot \|\vec{b}\|$.

3.1.2. Алгебраическое определение (в координатах)

Если векторы \vec{a} и \vec{b} заданы своими координатами в ортонормированном базисе (например, стандартном базисе в \mathbb{R}^n):

$$\vec{a} = (a_1, a_2, \dots, a_n)$$

$$\vec{b} = (b_1, b_2, \dots, b_n)$$

то их скалярное произведение равно сумме произведений соответствующих координат:

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + \dots + a_n b_n = \sum_{i=1}^n a_i b_i$$

Это определение наиболее удобно для вычислений.

В матричной записи, если \vec{a} и \vec{b} – векторы-столбцы, то $\vec{a} \cdot \vec{b} = \vec{a}^T \vec{b}$.

Связь между определениями: Равенство этих двух определений следует из теоремы косинусов.

3.2. Свойства Скалярного Произведения

1. Коммутативность (симметричность):

$$\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$$

2. Дистрибутивность относительно сложения векторов:

$$(\vec{a} + \vec{b}) \cdot \vec{c} = \vec{a} \cdot \vec{c} + \vec{b} \cdot \vec{c}$$

3. Ассоциативность относительно умножения на скаляр:

$$(k\vec{a}) \cdot \vec{b} = k(\vec{a} \cdot \vec{b}) = \vec{a} \cdot (k\vec{b})$$

где k – любой скаляр.

4. Неотрицательность скалярного квадрата:

$$\vec{a} \cdot \vec{a} = \|\vec{a}\|^2 \geq 0$$

Причем $\vec{a} \cdot \vec{a} = 0$ тогда и только тогда, когда $\vec{a} = \vec{0}$ (нулевой вектор).

Из этого свойства следует, что длина (норма) вектора может быть выражена через скалярное произведение: $||\vec{a}|| = \sqrt{\vec{a} \cdot \vec{a}}$.

5. Условие ортогональности векторов:

Ненулевые векторы \vec{a} и \vec{b} ортогональны (перпендикулярны) тогда и только тогда, когда их скалярное произведение равно нулю:

$$\vec{a} \cdot \vec{b} = 0 \iff \vec{a} \perp \vec{b}$$

3.3. Применение Скалярного Произведения

Скалярное произведение имеет множество применений в различных областях математики, физики и компьютерных наук, включая обучение с подкреплением.

1. Вычисление длины (нормы) вектора:

Как уже упоминалось, $||\vec{a}|| = \sqrt{\vec{a} \cdot \vec{a}}$.

2. Вычисление угла между векторами:

Из геометрического определения следует:

$$\cos(\theta) = \frac{\vec{a} \cdot \vec{b}}{||\vec{a}|| \cdot ||\vec{b}||}$$

Эта формула позволяет найти угол между двумя ненулевыми векторами, если известны их координаты.

3. Проекция одного вектора на другой:

Скалярная проекция вектора \vec{b} на направление вектора \vec{a} (если $\vec{a} \neq \vec{0}$) равна:

$$\text{proj}_a \vec{b} = \frac{\vec{a} \cdot \vec{b}}{||\vec{a}||} = ||\vec{b}|| \cos \theta$$

Векторная проекция \vec{b} на \vec{a} равна:

$$\vec{\text{proj}}_a \vec{b} = \left(\frac{\vec{a} \cdot \vec{b}}{||\vec{a}||^2} \right) \vec{a}$$

4. Работа силы в физике:

Если сила \vec{F} постоянна и тело перемещается по прямой на вектор \vec{s} , то работа W , совершаемая силой, равна $W = \vec{F} \cdot \vec{s}$.

Применение скалярного произведения в Обучении с Подкреплением (RL):

- **Измерение сходства между векторами:** Косинус угла, вычисляемый через скалярное произведение, часто используется как мера **косинусного сходства (cosine similarity)** между двумя векторами. Если векторы представляют, например, состояния или признаки, косинусное сходство может показать, насколько они "похожи" по направлению, независимо от их величины.

$$\text{similarity}(\vec{a}, \vec{b}) = \cos(\theta) = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \cdot \|\vec{b}\|}$$

Это используется в RL, например, для сравнения векторов признаков состояний, для кластеризации состояний или для поиска ближайших соседей.

- **Линейная аппроксимация функций ценности/стратегий:** В RL часто аппроксимируют функцию ценности $V(s)$ или Q-функцию $Q(s, a)$ как линейную комбинацию признаков $\vec{\phi}(s)$ с весами \vec{w} :

$$V(s) \approx \vec{w}^T \vec{\phi}(s) = \vec{w} \cdot \vec{\phi}(s)$$

Здесь скалярное произведение вектора весов на вектор признаков дает оценку ценности состояния. Обучение заключается в подборе таких весов \vec{w} , чтобы эта аппроксимация была как можно точнее.

- **Вычисление градиентов:** В алгоритмах, использующих градиентный спуск (например, в градиентах стратегий или при обучении нейронных сетей), скалярные произведения возникают при вычислении компонент градиента.
- **Ортогонализация признаков:** В некоторых случаях для улучшения стабильности и скорости обучения моделей признаки ортогонализуют (делают их скалярное произведение равным нулю). Процесс Грама-Шмидта, использующий скалярные произведения, позволяет построить ортонормированный базис из набора линейно независимых векторов.
- **Ядерные методы (Kernel Methods):** В более продвинутых техниках RL (например, Kernel-based RL) используются функции ядра $K(\vec{x}, \vec{y})$, которые часто можно представить как скалярное произведение векторов в некотором (возможно, бесконечномерном) пространстве признаков: $K(\vec{x}, \vec{y}) = \Phi(\vec{x}) \cdot \Phi(\vec{y})$. Это позволяет работать с нелинейными зависимостями.
- **Определение перпендикулярности/параллельности:** В задачах навигации или анализа траекторий агента скалярное произведение может помочь определить, движется ли агент перпендикулярно или параллельно некоторому направлению, или насколько его текущее действие соответствует целевому направлению.

Понимание скалярного произведения и его свойств является ключевым для работы с векторами в многомерных пространствах, что постоянно встречается в задачах обучения с подкреплением.

Раздел 4: Собственные Значения и Собственные Векторы. Диагонализация Матриц

4.1. Определение Собственных Значений и Собственных Векторов

Пусть A – квадратная матрица порядка n с действительными (или комплексными) элементами.

Ненулевой вектор-столбец $\vec{x} \in \mathbb{R}^n$ (или \mathbb{C}^n) называется **собственным вектором** матрицы A , если существует такое число λ (действительное или комплексное), что выполняется равенство:

$$A\vec{x} = \lambda\vec{x}$$

Число λ называется **собственным значением** (или **собственным числом**) матрицы A , соответствующим собственному вектору \vec{x} .

Геометрический смысл:

Если \vec{x} – собственный вектор матрицы A , то умножение матрицы A на вектор \vec{x} не изменяет направления вектора \vec{x} (или изменяет его на противоположное, если $\lambda < 0$). Происходит лишь масштабирование (растяжение или сжатие) вектора \vec{x} в λ раз. Линейное преобразование, задаваемое матрицей A , действует вдоль направления собственного вектора \vec{x} как простое умножение на скаляр λ .

- Если $\lambda > 1$, вектор \vec{x} растягивается.
- Если $0 < \lambda < 1$, вектор \vec{x} сжимается.
- Если $\lambda = 1$, вектор \vec{x} не изменяется (является неподвижной точкой преобразования, если не считать масштабирования).
- Если $\lambda < 0$, вектор \vec{x} меняет направление на противоположное и масштабируется в $|\lambda|$ раз.
- Если $\lambda = 0$, то $A\vec{x} = \vec{0}$. Это означает, что собственный вектор \vec{x} принадлежит ядру (нуль-пространству) матрицы A . Собственное значение $\lambda = 0$ существует тогда и только тогда, когда матрица A вырождена ($\det(A) = 0$).

Важные замечания:

- Собственный вектор по определению должен быть **ненулевым**. Нулевой вектор $\vec{0}$ всегда удовлетворяет уравнению $A\vec{0} = \lambda\vec{0}$ для любого λ , но не несет полезной информации.
- Если \vec{x} – собственный вектор, соответствующий собственному значению λ , то любой коллинеарный ему ненулевой вектор $k\vec{x}$ (где $k \neq 0$ – скаляр) также является собственным вектором, соответствующим тому же собственному значению λ , так как $A(k\vec{x}) = k(A\vec{x}) = k(\lambda\vec{x}) = \lambda(k\vec{x})$.
- Одному собственному значению может соответствовать несколько линейно независимых собственных векторов.

4.2. Нахождение Собственных Значений и Собственных Векторов

4.2.1. Характеристическое Уравнение

Перепишем основное уравнение $A\vec{x} = \lambda\vec{x}$:

$$A\vec{x} - \lambda\vec{x} = \vec{0}$$

$$A\vec{x} - \lambda I\vec{x} = \vec{0}$$

$$(A - \lambda I)\vec{x} = \vec{0}$$

где I – единичная матрица того же порядка, что и A .

Это однородная система линейных алгебраических уравнений (СЛАУ) относительно компонент вектора \vec{x} . Поскольку мы ищем ненулевые решения \vec{x} для этой системы, ее матрица коэффициентов $(A - \lambda I)$ должна быть вырожденной. Это означает, что определитель этой матрицы должен быть равен нулю:

$$\det(A - \lambda I) = 0$$

Это уравнение называется **характеристическим уравнением** матрицы A . Его корни $\lambda_1, \lambda_2, \dots, \lambda_n$ являются **собственными значениями** матрицы A .

Для матрицы $n \times n$ характеристическое уравнение представляет собой многочлен n -й степени относительно λ , называемый **характеристическим многочленом** $P(\lambda) = \det(A - \lambda I)$. Согласно основной теореме алгебры, такой многочлен имеет ровно n корней (с учетом их кратности) в поле комплексных чисел. Если матрица A действительная, то комплексные собственные значения (если они есть) появляются сопряженными парами.

Пример для матрицы 2x2:

Пусть $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$.

Тогда $A - \lambda I = \begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{pmatrix}$.

Характеристическое уравнение: $\det(A - \lambda I) = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = 0$.

Это квадратное уравнение относительно λ : $\lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0$.

Заметим, что $a_{11} + a_{22} = \text{tr}(A)$ (след матрицы), а $a_{11}a_{22} - a_{12}a_{21} = \det(A)$.

Таким образом, для матрицы 2x2: $\lambda^2 - \text{tr}(A)\lambda + \det(A) = 0$.

4.2.2. Нахождение Собственных Векторов

После того как найдены собственные значения λ_i , для каждого из них необходимо найти соответствующие собственные векторы.

Для каждого конкретного собственного значения λ_k подставляем его в систему $(A - \lambda_k I)\vec{x} = \vec{0}$ и решаем эту однородную СЛАУ относительно компонент вектора $\vec{x} = (x_1, x_2, \dots, x_n)^T$.

Поскольку λ_k является корнем характеристического уравнения, определитель матрицы $(A - \lambda_k I)$ равен нулю. Это означает, что система будет иметь бесконечно много ненулевых решений (строки матрицы линейно зависимы). Решения образуют **собственное подпространство** E_{λ_k} , соответствующее собственному значению λ_k .

Любой ненулевой вектор из этого подпространства является собственным вектором для λ_k .

Обычно находят **базис** этого собственного подпространства. Размерность собственного подпространства E_{λ_k} называется **геометрической кратностью** собственного значения λ_k . Геометрическая кратность не превышает алгебраической кратности (кратности корня λ_k в характеристическом многочлене).

Алгоритм нахождения собственных значений и векторов:

1. Для данной квадратной матрицы A составить матрицу $(A - \lambda I)$.
2. Вычислить определитель $\det(A - \lambda I)$ и приравнять его к нулю. Это даст характеристическое уравнение.
3. Решить характеристическое уравнение относительно λ , чтобы найти собственные значения $\lambda_1, \lambda_2, \dots, \lambda_n$.
4. Для каждого найденного собственного значения λ_k :
 - а. Подставить λ_k в систему уравнений $(A - \lambda_k I)\vec{x} = \vec{0}$.
 - б. Решить эту однородную СЛАУ, чтобы найти общее решение для \vec{x} . Это решение опишет все собственные векторы, соответствующие λ_k . Обычно представляют один или несколько линейно независимых собственных векторов (базис собственного подпространства).

4.3. Свойства Собственных Значений и Собственных Векторов

1. **Сумма собственных значений** равна следу матрицы: $\sum_{i=1}^n \lambda_i = \text{tr}(A) = \sum_{i=1}^n a_{ii}$.
2. **Произведение собственных значений** равно определителю матрицы:
$$\prod_{i=1}^n \lambda_i = \det(A).$$
3. Собственные значения **треугольной** (в частности, диагональной) матрицы равны ее диагональным элементам.
4. Если матрица A невырождена ($\det(A) \neq 0$), то все ее собственные значения ненулевые. Если A вырождена, то хотя бы одно собственное значение равно нулю.
5. Если λ – собственное значение матрицы A с собственным вектором \vec{x} , то:
 - $k\lambda$ – собственное значение матрицы kA с тем же собственным вектором \vec{x} (для скаляра k).
 - λ^m – собственное значение матрицы A^m (для целого $m > 0$) с тем же собственным вектором \vec{x} .
 - Если A обратима, то $1/\lambda$ – собственное значение матрицы A^{-1} с тем же собственным вектором \vec{x} . (Заметим, что $\lambda \neq 0$, так как A обратима).
6. Собственные векторы, соответствующие **различным** собственным значениям, **линейно независимы**.
7. Для **симметрической действительной матрицы** A ($A = A^T$):
 - Все собственные значения действительны.

- Собственные векторы, соответствующие различным собственным значениям, ортогональны (их скалярное произведение равно нулю).
- Существует ортонормированный базис из собственных векторов (т.е. матрицу можно диагонализировать ортогональным преобразованием).

4.4. Диагонализация Матриц

Квадратная матрица A порядка n называется **диагонализируемой**, если она подобна некоторой диагональной матрице D . То есть, если существует такая невырожденная матрица P , что:

$$P^{-1}AP = D$$

где $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ – диагональная матрица, на диагонали которой стоят собственные значения матрицы A .

Условие диагонализируемости:

Матрица A порядка n диагонализируема тогда и только тогда, когда она имеет n линейно независимых собственных векторов.

- Если все n собственных значений матрицы A различны, то матрица A всегда диагонализируема (так как n собственных векторов, соответствующих различным собственным значениям, всегда линейно независимы).
- Если среди собственных значений есть кратные, то матрица диагонализируема тогда и только тогда, когда для каждого собственного значения его геометрическая кратность (размерность соответствующего собственного подпространства) равна его алгебраической кратности (кратности как корня характеристического многочлена).

Процесс диагонализации:

1. Найти все собственные значения $\lambda_1, \dots, \lambda_n$ матрицы A .
2. Для каждого собственного значения найти соответствующий линейно независимый собственный вектор (или базис собственного подпространства, если кратность > 1).
3. Если удалось найти n линейно независимых собственных векторов $\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n$, то матрица A диагонализируема.
4. Составить матрицу P , столбцами которой являются эти n линейно независимых собственных векторов: $P = [\vec{p}_1 | \vec{p}_2 | \dots | \vec{p}_n]$.
5. Тогда $P^{-1}AP = D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, где λ_i – собственное значение, соответствующее собственному вектору \vec{p}_i . Порядок собственных значений в D соответствует порядку собственных векторов-столбцов в P .

Частный случай: Симметрические матрицы

Любая действительная симметрическая матрица A ($A = A^T$) диагонализируема. Более того, ее можно диагонализировать с помощью **ортогональной матрицы** P (т.е. $P^{-1} = P^T$).

). В этом случае столбцы матрицы P образуют ортонормированный базис из собственных векторов матрицы A .

$$P^T A P = D$$

Это утверждение известно как **спектральная теорема** для симметрических матриц.

4.5. Применение Собственных Значений и Векторов в Обучении с Подкреплением (RL)

Собственные значения и собственные векторы играют фундаментальную роль во многих аспектах математики и инженерии, и их концепции находят применение в RL:

- **Анализ матриц переходных вероятностей:** В Марковских процессах принятия решений (MDP) матрица переходных вероятностей P описывает динамику системы. Собственные значения этой матрицы могут дать информацию о поведении системы в долгосрочной перспективе. Например, для эргодических цепей Маркова наибольшее по модулю собственное значение равно 1, и соответствующий ему собственный вектор связан со стационарным распределением состояний.
- **Сходимость итерационных алгоритмов:** Многие алгоритмы в RL, такие как Value Iteration или Policy Iteration, являются итерационными. Скорость их сходимости может быть связана с собственными значениями операторов, лежащих в основе этих итераций (например, оператора Беллмана). Фактор дисконтирования γ часто обеспечивает сжатие оператора, что гарантирует сходимость, и это связано со спектральным радиусом (максимальным модулем собственного значения) соответствующей матрицы или оператора.
- **Анализ главных компонент (PCA):** PCA – это метод понижения размерности, который основан на нахождении собственных векторов ковариационной матрицы данных. Собственные векторы (главные компоненты) указывают направления наибольшей дисперсии в данных, а соответствующие собственные значения показывают величину этой дисперсии. В RL PCA может использоваться для предобработки признаков состояния, уменьшения размерности пространства состояний или действий.
- **Устойчивость систем управления:** В теории управления, тесно связанной с RL, собственные значения матрицы динамики системы определяют ее устойчивость. Если все собственные значения имеют отрицательную действительную часть (для непрерывных систем) или модуль меньше единицы (для дискретных систем), система устойчива.
- **Разложения матриц:** Диагонализация является формой матричного разложения. Другие разложения, такие как Сингулярное Разложение Значений (SVD), тесно связаны с собственными значениями и векторами (SVD использует собственные значения матриц $A^T A$ и $A A^T$). SVD широко используется в RL для аппроксимации

матриц, понижения размерности и в рекомендательных системах (которые могут рассматриваться как частный случай RL).

- **Квадратичные формы и оптимизация:** Собственные значения симметрической матрицы определяют характер квадратичной формы $x^T A x$. Если все $\lambda_i > 0$, форма положительно определена (имеет минимум в нуле). Это важно в задачах оптимизации, которые часто возникают в RL.
- **Графовые методы в RL:** Если состояния и переходы в MDP представляются графом, то спектральные свойства матрицы смежности или лапласиана графа (которые связаны с их собственными значениями и векторами) могут использоваться для анализа структуры пространства состояний, кластеризации состояний или построения признаков (например, Laplacian Eigenmaps).

Понимание концепции собственных значений и векторов позволяет глубже анализировать поведение линейных систем и операторов, что является важным для разработки и понимания многих алгоритмов обучения с подкреплением.

Раздел 5: Сингулярное Разложение (SVD)

5.1. Определение Сингулярного Разложения

Сингулярное разложение (Singular Value Decomposition, SVD) – это фундаментальное разложение прямоугольной (или квадратной) матрицы на произведение трех матриц особого вида. Оно является обобщением понятия диагонализации для произвольных матриц и имеет широкое применение в линейной алгебре, статистике, машинном обучении и, в частности, в обучении с подкреплением.

Для любой действительной матрицы A размера $m \times n$ (где m может быть не равно n) существует сингулярное разложение вида:

$$A = U \Sigma V^T$$

где:

- U – ортогональная матрица размера $m \times m$. Столбцы U называются **левыми сингулярными векторами**.
(Ортогональная матрица означает, что $U^T U = U U^T = I_m$, где I_m – единичная матрица $m \times m$.)
- Σ – диагональная матрица размера $m \times n$ с неотрицательными действительными числами на главной диагонали, расположенными в порядке убывания. Эти числа $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ (где $r = \text{rank}(A)$ – ранг матрицы A) называются **сингулярными числами** (или сингулярными значениями) матрицы A . Остальные элементы Σ (включая диагональные элементы $\sigma_{r+1}, \dots, \sigma_{\min(m,n)}$ и все недиагональные элементы) равны нулю.
Форма матрицы Σ зависит от соотношения m и n :
 - Если $m = n$ (квадратная матрица), $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_n)$.

- Если $m > n$ (высокая матрица), $\Sigma = \begin{pmatrix} D \\ 0 \end{pmatrix}$, где $D = \text{diag}(\sigma_1, \dots, \sigma_n)$ – матрица $n \times n$.
- Если $m < n$ (широкая матрица), $\Sigma = (D \ 0)$, где $D = \text{diag}(\sigma_1, \dots, \sigma_m)$ – матрица $m \times m$.
- V – ортогональная матрица размера $n \times n$. Столбцы V называются **правыми сингулярными векторами**.
(Ортогональная матрица означает, что $V^T V = V V^T = I_n$, где I_n – единичная матрица $n \times n$.)

Связь с собственными значениями:

Сингулярные числа σ_i матрицы A являются квадратными корнями из собственных значений матриц $A^T A$ (и AA^T).

- Столбцы V (правые сингулярные векторы) являются собственными векторами матрицы $A^T A$.

$$A^T A = (U \Sigma V^T)^T (U \Sigma V^T) = V \Sigma^T U^T U \Sigma V^T = V (\Sigma^T \Sigma) V^T.$$
 Так как V ортогональна, $V^{-1} = V^T$. Тогда $(A^T A)V = V(\Sigma^T \Sigma)$. Это показывает, что столбцы V – собственные векторы $A^T A$, а диагональные элементы $\Sigma^T \Sigma$ (которые равны σ_i^2) – соответствующие собственные значения.
- Столбцы U (левые сингулярные векторы) являются собственными векторами матрицы AA^T .

$$AA^T = (U \Sigma V^T)(U \Sigma V^T)^T = U \Sigma V^T V \Sigma^T U^T = U (\Sigma \Sigma^T) U^T.$$
 Аналогично, $(AA^T)U = U(\Sigma \Sigma^T)$. Столбцы U – собственные векторы AA^T , а диагональные элементы $\Sigma \Sigma^T$ (которые также равны σ_i^2) – соответствующие собственные значения.

5.2. Геометрический Смысл SVD

SVD можно интерпретировать как разложение линейного преобразования, задаваемого матрицей A , на три последовательных геометрических преобразования:

1. **Вращение или отражение в пространстве исходных векторов (задаваемое V^T):** Матрица V^T преобразует исходные векторы. Поскольку V ортогональна, V^T представляет собой вращение (возможно, с отражением).
2. **Масштабирование вдоль новых осей (задаваемое Σ):** Матрица Σ растягивает или сжимает пространство вдоль канонических осей. Коэффициенты масштабирования – это сингулярные числа σ_i . Если какое-то $\sigma_i = 0$, то преобразование "схлопывает" пространство вдоль соответствующей оси (понижает размерность).
3. **Вращение или отражение в пространстве результирующих векторов (задаваемое U):** Матрица U выполняет вращение (возможно, с отражением) в пространстве, куда отображаются векторы.

Таким образом, любое линейное преобразование можно представить как поворот, масштабирование по осям и еще один поворот.

5.3. Свойства Сингулярного Разложения

1. **Существование и уникальность:** SVD существует для любой матрицы A .

Сингулярные числа σ_i определены однозначно. Сингулярные векторы (столбцы U и V) определены однозначно с точностью до знака, если все сингулярные числа различны. Если есть кратные сингулярные числа, то соответствующие сингулярные векторы образуют подпространства, и любой ортонормированный базис в этих подпространствах может быть выбран.

2. **Ранг матрицы:** Ранг матрицы A равен количеству ее ненулевых сингулярных чисел r .

3. **Нормы матриц:**

- **Фробениусова норма:** $\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} = \sqrt{\sum_{k=1}^r \sigma_k^2}$.
- **Спектральная норма (или норма L_2):** $\|A\|_2 = \max_{\vec{x} \neq \vec{0}} \frac{\|A\vec{x}\|_2}{\|\vec{x}\|_2} = \sigma_1$ (наибольшее сингулярное число).

4. **Определитель (для квадратных матриц):** Если A – квадратная матрица $n \times n$, то $|\det(A)| = \prod_{i=1}^n \sigma_i$.

5. **Обратная и псевдообратная матрицы:**

- Если A – квадратная невырожденная матрица, то $A^{-1} = V\Sigma^{-1}U^T$, где $\Sigma^{-1} = \text{diag}(1/\sigma_1, \dots, 1/\sigma_n)$.
- **Псевдообратная матрица Мура-Пенроуза (A^+):** Для любой матрицы A (включая прямоугольные и вырожденные) псевдообратная матрица определяется как $A^+ = V\Sigma^+U^T$. Матрица Σ^+ получается из Σ путем транспонирования и замены каждого ненулевого диагонального элемента σ_i на $1/\sigma_i$. Псевдообратная матрица используется для нахождения решения системы $A\vec{x} = \vec{b}$ с минимальной нормой L_2 среди всех решений, минимизирующих $\|A\vec{x} - \vec{b}\|_2$ (решение задачи наименьших квадратов).

6. **Аппроксимация матриц низким рангом (Теорема Эккарта-Янга):** SVD позволяет найти наилучшее приближение матрицы A матрицей A_k заданного ранга $k < r$. Это делается путем обнуления всех сингулярных чисел, кроме первых k наибольших: $A_k = U\Sigma_k V^T$, где Σ_k – это матрица Σ , в которой $\sigma_{k+1}, \dots, \sigma_r$ заменены нулями. Матрица A_k является наилучшим приближением A ранга k в смысле фробениусовой нормы и спектральной нормы: $\|A - A_k\|_F$ и $\|A - A_k\|_2$ минимальны.

Это свойство лежит в основе метода главных компонент (PCA) и многих методов понижения размерности.

5.4. Усеченное (Экономное) SVD

Если $m > n$ или $m < n$, то некоторые части матриц U , Σ , V могут быть избыточными (содержать нулевые блоки или соответствовать нулевым сингулярным числам).

Экономное (thin) SVD:

Если $m \geq n$, то $A = U_n \Sigma_n V^T$, где:

- U_n – матрица $m \times n$, содержащая первые n столбцов U .
- Σ_n – квадратная диагональная матрица $n \times n$, содержащая первые n сингулярных чисел $\sigma_1, \dots, \sigma_n$.
- V – матрица $n \times n$ (та же, что и в полном SVD).

Если $m < n$, то $A = U \Sigma_m V_m^T$, где:

- U – матрица $m \times m$ (та же, что и в полном SVD).
- Σ_m – квадратная диагональная матрица $m \times m$, содержащая первые m сингулярных чисел $\sigma_1, \dots, \sigma_m$.
- V_m – матрица $n \times m$, содержащая первые m столбцов V .

Компактное (compact) SVD (или SVD с рангом r):

$A = U_r \Sigma_r V_r^T$, где $r = \text{rank}(A)$.

- U_r – матрица $m \times r$, содержащая первые r столбцов U (соответствующие ненулевым сингулярным числам).
 - Σ_r – квадратная диагональная матрица $r \times r$, содержащая r ненулевых сингулярных чисел $\sigma_1, \dots, \sigma_r$.
 - V_r – матрица $n \times r$, содержащая первые r столбцов V .
- Это представление наиболее компактно и содержит всю информацию о ненулевых компонентах преобразования.

5.5. Применение SVD в Обучении с Подкреплением (RL)

SVD является мощным инструментом анализа и обработки данных, который находит применение в различных аспектах RL:

1. **Понижение размерности пространства состояний/действий:** В задачах с большим или непрерывным пространством состояний/действий SVD (часто через PCA, который тесно связан с SVD ковариационной матрицы) может использоваться для нахождения низкоразмерного представления, сохраняющего наиболее важную информацию. Это помогает бороться с "проклятием размерности".
2. **Аппроксимация функций ценности или стратегий:** Если функция ценности (value function) или стратегия (policy) представляется в виде матрицы (например, Q-таблица в дискретных случаях), SVD может использоваться для ее аппроксимации матрицей меньшего ранга. Это может быть полезно для сжатия модели, ускорения вычислений или сглаживания.

3. Решение систем линейных уравнений и задачи наименьших квадратов:

Многие алгоритмы RL, особенно те, что основаны на аппроксимации функций (например, LSTD - Least Squares Temporal Difference), сводятся к решению систем линейных уравнений или задач наименьших квадратов. SVD и псевдообратная матрица предоставляют робастные методы для их решения, особенно когда матрицы плохо обусловлены или вырождены.

4. Анализ матриц переходных вероятностей или матриц вознаграждений: SVD

может применяться для анализа структуры этих матриц, выявления скрытых зависимостей или построения их низкоранговых аппроксимаций.

5. Рекомендательные системы как RL-задачи: В некоторых формулировках

рекомендательные системы можно рассматривать как задачи RL. SVD является классическим методом для матричной факторизации в рекомендательных системах (например, для предсказания рейтингов), где матрица "пользователь-элемент" разлагается для выявления скрытых факторов.

6. Инициализация весов в нейронных сетях: Иногда техники, связанные с SVD,

используются для анализа или инициализации весовых матриц в глубоких нейронных сетях, применяемых в RL, для улучшения стабильности обучения.

7. Робастность к шуму: Низкоранговая аппроксимация с помощью SVD может

помочь отфильтровать шум в данных, так как сингулярные векторы, соответствующие малым сингулярным числам, часто связаны с шумовыми компонентами.

Понимание SVD дает возможность применять мощные методы линейной алгебры для анализа, упрощения и решения сложных задач, возникающих в теории и практике обучения с подкреплением.

Раздел 6: Дополнительные Темы Линейной Алгебры

6.1. Квадратичные Формы

6.1.1. Определение Квадратичной Формы

Квадратичной формой от n переменных x_1, x_2, \dots, x_n называется однородный многочлен второй степени относительно этих переменных. Каждое слагаемое такого многочлена содержит либо квадрат одной из переменных (x_i^2), либо произведение двух различных переменных ($x_i x_j$).

Общий вид квадратичной формы $Q(x_1, x_2, \dots, x_n)$:

$$Q(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$$

где a_{ij} — действительные коэффициенты.

Например, для двух переменных x_1, x_2 :

$$Q(x_1, x_2) = a_{11}x_1^2 + a_{12}x_1x_2 + a_{21}x_2x_1 + a_{22}x_2^2$$

Поскольку $x_1x_2 = x_2x_1$, мы можем объединить коэффициенты при x_1x_2 и x_2x_1 . Обычно квадратичную форму записывают так, чтобы матрица коэффициентов была симметричной. Для этого коэффициент при x_ix_j (где $i \neq j$) делят пополам между a_{ij} и a_{ji} .

Тогда для двух переменных:

$$Q(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$$

(здесь a_{12} в формуле соответствует a_{12} из симметричной матрицы, а исходный коэффициент при x_1x_2 был бы $a_{12} + a_{21}$, который стал $2a_{12}$ при $a_{12} = a_{21}$).

Для трех переменных x_1, x_2, x_3 :

$$Q(x_1, x_2, x_3) = a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3$$

6.1.2. Матричная Запись Квадратичной Формы

Любую квадратичную форму можно представить в матричном виде:

$$Q(\vec{x}) = \vec{x}^T A \vec{x}$$

где:

- $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ – вектор-столбец переменных.
- $\vec{x}^T = (x_1, x_2, \dots, x_n)$ – вектор-строка переменных (транспонированный \vec{x}).
- A – **матрица квадратичной формы**. Это симметрическая матрица $n \times n$, элементы которой a_{ij} являются коэффициентами квадратичной формы. На главной диагонали A стоят коэффициенты при квадратах x_i^2 (т.е. a_{ii}), а вне диагонали элементы $a_{ij} = a_{ji}$ равны половине коэффициента при произведении x_ix_j (для $i \neq j$).

Пример:

Для $Q(x_1, x_2) = 3x_1^2 - 4x_1x_2 + 2x_2^2$.

Матрица $A = \begin{pmatrix} 3 & -2 \\ -2 & 2 \end{pmatrix}$.

Проверка:

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 3 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 3x_1 - 2x_2 \\ -2x_1 + 2x_2 \end{pmatrix} = x_1(3x_1 - 2x_2) + x_2(-2x_1 + 2x_2) = 3x_1^2 - 4x_1x_2 + 2x_2^2$$

Дискриминант квадратичной формы – это определитель ее матрицы $\det(A)$.

Ранг квадратичной формы – это ранг ее матрицы $\text{rank}(A)$.

6.1.3. Приведение Квадратичной Формы к Каноническому Виду

Канонический вид квадратичной формы – это представление, в котором отсутствуют слагаемые с произведениями различных переменных ($x_i x_j$ для $i \neq j$), т.е. форма содержит только квадраты переменных:

$$Q(y_1, \dots, y_n) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2$$

где y_1, \dots, y_n – новые переменные, полученные из x_1, \dots, x_n с помощью некоторого линейного невырожденного преобразования $\vec{x} = P\vec{y}$, а λ_i – действительные коэффициенты.

Матрица квадратичной формы в каноническом виде является диагональной:

$$D = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Существует несколько методов приведения квадратичной формы к каноническому виду:

1. Метод Лагранжа (метод выделения полных квадратов):

Этот метод заключается в последовательном выделении полных квадратов. Он всегда позволяет привести форму к каноническому виду.

- Если есть хотя бы один коэффициент $a_{ii} \neq 0$ (например, $a_{11} \neq 0$), собирают все слагаемые, содержащие x_1 , и дополняют их до полного квадрата.
- После этого делают замену переменных, чтобы избавиться от x_1 в оставшейся части формы.
- Процесс повторяют для оставшихся переменных.
- Если все $a_{ii} = 0$, но есть $a_{ij} \neq 0$ (например $a_{12} \neq 0$), делают предварительную невырожденную замену (например, $x_1 = y_1 + y_2, x_2 = y_1 - y_2, x_k = y_k$ для $k > 2$), чтобы появился член с квадратом.

2. Метод ортогональных преобразований (приведение к главным осям):

Этот метод использует свойства симметрических матриц. Матрицу A квадратичной формы можно диагонализировать с помощью ортогональной матрицы P (столбцы которой – ортонормированные собственные векторы A):

$$P^T A P = D = \text{diag}(\lambda_1, \dots, \lambda_n)$$

где λ_i – собственные значения матрицы A .

Если сделать замену переменных $\vec{x} = P\vec{y}$, то квадратичная форма примет вид:

$$Q(\vec{x}) = \vec{x}^T A \vec{x} = (P\vec{y})^T A (P\vec{y}) = \vec{y}^T (P^T A P) \vec{y} = \vec{y}^T D \vec{y} = \sum_{i=1}^n \lambda_i y_i^2$$

Коэффициенты λ_i в каноническом виде, полученном ортогональным преобразованием, являются собственными значениями матрицы A .

6.1.4. Закон Инерции Квадратичных Форм

Закон инерции гласит, что число положительных, отрицательных и нулевых коэффициентов в каноническом виде квадратичной формы не зависит от способа приведения к этому виду (т.е. от выбора невырожденного линейного преобразования).

- Число положительных коэффициентов называется **положительным индексом инерции** (r_+).
- Число отрицательных коэффициентов называется **отрицательным индексом инерции** (r_-).
- Ранг квадратичной формы $r = r_+ + r_-$.
- **Сигнатура** квадратичной формы – это пара (r_+, r_-) или разность $s = r_+ - r_-$.

6.1.5. Знакоопределенность Квадратичных Форм

В зависимости от знаков коэффициентов λ_i в каноническом виде (или, что эквивалентно, от знаков собственных значений матрицы A), квадратичная форма $Q(\vec{x})$ называется:

- **Положительно определенной**, если $Q(\vec{x}) > 0$ для всех ненулевых векторов \vec{x} . Это происходит, когда все $\lambda_i > 0$ (все r_n собственных значений положительны).
- **Отрицательно определенной**, если $Q(\vec{x}) < 0$ для всех ненулевых векторов \vec{x} . Это происходит, когда все $\lambda_i < 0$ (все r_n собственных значений отрицательны).
- **Положительно полуопределенной**, если $Q(\vec{x}) \geq 0$ для всех \vec{x} (и существует $\vec{x} \neq 0$, для которого $Q(\vec{x}) = 0$). Это происходит, когда все $\lambda_i \geq 0$ и хотя бы одно $\lambda_i = 0$.
- **Отрицательно полуопределенной**, если $Q(\vec{x}) \leq 0$ для всех \vec{x} (и существует $\vec{x} \neq 0$, для которого $Q(\vec{x}) = 0$). Это происходит, когда все $\lambda_i \leq 0$ и хотя бы одно $\lambda_i = 0$.
- **Неопределенной (знакопеременной)**, если $Q(\vec{x})$ принимает как положительные, так и отрицательные значения. Это происходит, когда есть и положительные, и отрицательные λ_i .

Критерий Сильвестра (для определения знакоопределенности без приведения к каноническому виду):

Пусть A – матрица квадратичной формы. Рассмотрим ее **главные угловые миноры**:

$$\Delta_1 = a_{11}$$

$$\Delta_2 = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$\Delta_3 = \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

...

$$\Delta_n = \det(A)$$

- Квадратичная форма **положительно определена** тогда и только тогда, когда все ее главные угловые миноры строго положительны: $\Delta_k > 0$ для всех $k = 1, \dots, n$.

- Квадратичная форма **отрицательно определена** тогда и только тогда, когда знаки ее главных угловых миноров чередуются, начиная с минуса:

$$\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \dots, (-1)^n \Delta_n > 0.$$

Для полуопределенности критерий Сильвестра усложняется (требуется рассмотрение всех главных миноров, а не только угловых).

6.1.6. Применение Квадратичных Форм в Обучении с Подкреплением

- **Функции Ляпунова и анализ устойчивости:** В теории управления и динамических системах (которые тесно связаны с RL) квадратичные формы используются для построения функций Ляпунова для анализа устойчивости систем. Положительно определенная функция Ляпунова, производная которой отрицательно определена вдоль траекторий системы, гарантирует асимптотическую устойчивость.
- **Оптимизационные задачи:** Многие задачи оптимизации в RL (например, при поиске оптимальной стратегии или аппроксимации функций ценности) могут включать минимизацию или максимизацию функций, которые локально могут быть аппроксимированы квадратичными формами (например, в методе Ньютона или методах второго порядка). Знакоопределенность гессиана (матрицы вторых производных), который определяет локальную квадратичную аппроксимацию, указывает на наличие минимума, максимума или седловой точки.
- **Метрики и расстояния:** Квадратичные формы могут определять метрики или расстояния в пространстве состояний или параметров. Например, расстояние Махаланобиса $d^2(\vec{x}, \vec{\mu}) = (\vec{x} - \vec{\mu})^T \Sigma^{-1} (\vec{x} - \vec{\mu})$ является квадратичной формой и используется, когда компоненты вектора коррелированы.
- **Аппроксимация функций:** В некоторых методах аппроксимации функций (например, в методе наименьших квадратов) целевая функция ошибки является квадратичной формой от параметров аппроксимирующей функции.
- **Регуляризация:** В задачах обучения, включая RL, часто добавляют регуляризационные члены к функции потерь, чтобы предотвратить переобучение. L2-регуляризация (например, $\lambda \|\vec{w}\|^2$) является квадратичной формой от весов модели.

Понимание свойств квадратичных форм важно для анализа поведения и сходимости многих алгоритмов, а также для формулирования и решения оптимизационных задач в RL.

6.2. Проекции и Ортогональные Дополнения

6.2.1. Ортогональное Дополнение

Пусть V – евклидово (или эрмитово) пространство, и W – его подпространство ($W \subset V$).

Ортогональным дополнением к подпространству W называется множество W^\perp (читается " W ортогональное" или " W перп"), состоящее из всех векторов в V , которые ортогональны каждому вектору из W . Формально:

$$W^\perp = \{\vec{v} \in V \mid (\vec{v}, \vec{w}) = 0 \text{ для всех } \vec{w} \in W\}$$

где (\vec{v}, \vec{w}) обозначает скалярное произведение векторов \vec{v} и \vec{w} .

Свойства ортогонального дополнения:

1. W^\perp также является подпространством пространства V .
2. Пересечение W и W^\perp содержит только нулевой вектор: $W \cap W^\perp = \{\vec{0}\}$.
3. Все пространство V может быть представлено как прямая сумма подпространства W и его ортогонального дополнения W^\perp : $V = W \oplus W^\perp$. Это означает, что любой вектор $\vec{v} \in V$ может быть единственным образом представлен в виде суммы $\vec{v} = \vec{w} + \vec{w}^\perp$, где $\vec{w} \in W$ и $\vec{w}^\perp \in W^\perp$.
4. $(W^\perp)^\perp = W$. Ортогональное дополнение к ортогональному дополнению есть исходное подпространство.
5. Если $\dim(V) = n$ и $\dim(W) = k$, то $\dim(W^\perp) = n - k$.

6.2.2. Ортогональная Проекция и Ортогональная Составляющая

Исходя из разложения $V = W \oplus W^\perp$, любой вектор $\vec{v} \in V$ можно однозначно записать как:

$$\vec{v} = \text{proj}_W \vec{v} + \text{ort}_W \vec{v}$$

где:

- $\text{proj}_W \vec{v}$ – **ортогональная проекция** вектора \vec{v} на подпространство W . Это компонента \vec{w} из разложения $\vec{v} = \vec{w} + \vec{w}^\perp$, где $\vec{w} \in W$.
- $\text{ort}_W \vec{v}$ – **ортогональная составляющая** вектора \vec{v} относительно подпространства W . Это компонента \vec{w}^\perp из разложения, где $\vec{w}^\perp \in W^\perp$. Ясно, что $\text{ort}_W \vec{v} = \vec{v} - \text{proj}_W \vec{v}$. Также $\text{ort}_W \vec{v} = \text{proj}_{W^\perp} \vec{v}$.

Свойства ортогональной проекции:

1. **Ближайшая точка:** Проекция $\text{proj}_W \vec{v}$ является вектором в W , который находится на наименьшем расстоянии от \vec{v} . То есть, для любого $\vec{u} \in W$, $\|\vec{v} - \text{proj}_W \vec{v}\| \leq \|\vec{v} - \vec{u}\|$. Равенство достигается только при $\vec{u} = \text{proj}_W \vec{v}$.
2. **Линейность:** Оператор проектирования $\text{proj}_W : V \rightarrow W$ является линейным оператором.

3. **Идемпотентность:** $\text{proj}_W(\text{proj}_W \vec{v}) = \text{proj}_W \vec{v}$. Повторное применение проекции не изменяет результат.

Нахождение ортогональной проекции:

- **Проекция на вектор (одномерное подпространство):**

Если $W = \text{span}\{\vec{a}\}$ – подпространство, натянутое на один ненулевой вектор \vec{a} , то проекция вектора \vec{v} на W (или на направление вектора \vec{a}) вычисляется по формуле:

$$\text{proj}_{\vec{a}} \vec{v} = \frac{(\vec{v}, \vec{a})}{(\vec{a}, \vec{a})} \vec{a} = \frac{(\vec{v}, \vec{a})}{\|\vec{a}\|^2} \vec{a}$$

- **Проекция на подпространство с ортонормированным базисом:**

Если $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k\}$ – ортонормированный базис подпространства W (т.е. $(\vec{u}_i, \vec{u}_j) = 0$ при $i \neq j$ и $(\vec{u}_i, \vec{u}_i) = \|\vec{u}_i\|^2 = 1$), то проекция вектора \vec{v} на W равна сумме проекций \vec{v} на каждый базисный вектор:

$$\text{proj}_W \vec{v} = (\vec{v}, \vec{u}_1) \vec{u}_1 + (\vec{v}, \vec{u}_2) \vec{u}_2 + \dots + (\vec{v}, \vec{u}_k) \vec{u}_k = \sum_{i=1}^k (\vec{v}, \vec{u}_i) \vec{u}_i$$

- **Проекция на подпространство с произвольным базисом (через матрицу проекции):**

Пусть столбцы матрицы A (размера $m \times k$) образуют базис подпространства $W \subset \mathbb{R}^m$. Тогда вектор $\text{proj}_W \vec{v}$ можно найти как $\vec{p} = A \hat{x}$, где \hat{x} – решение системы нормальных уравнений $A^T A \hat{x} = A^T \vec{v}$.

Если столбцы A линейно независимы (что так и есть, раз они базис), то матрица $A^T A$ обратима, и $\hat{x} = (A^T A)^{-1} A^T \vec{v}$.

Тогда проекция:

$$\text{proj}_W \vec{v} = A(A^T A)^{-1} A^T \vec{v}$$

Матрица $P = A(A^T A)^{-1} A^T$ называется **матрицей проекции** на подпространство W (которое является пространством столбцов A , $\text{Col}(A)$).

Свойства матрицы проекции P :

1. $P^T = P$ (симметрична).
2. $P^2 = P$ (идемпотентна).

6.2.3. Угол между Вектором и Подпространством

Углом между ненулевым вектором $\vec{v} \in V$ и подпространством $W \subset V$ называется наименьший угол между \vec{v} и произвольным ненулевым вектором $\vec{w} \in W$. Этот угол достигается, когда $\vec{w} = \text{proj}_W \vec{v}$ (если проекция не нулевая).

$$\angle(\vec{v}, W) = \angle(\vec{v}, \text{proj}_W \vec{v})$$

Косинус этого угла можно найти как:

$$\cos(\angle(\vec{v}, W)) = \frac{||\text{proj}_W \vec{v}||}{||\vec{v}||}$$

(при условии $\text{proj}_W \vec{v} \neq \vec{0}$ и $\vec{v} \neq \vec{0}$). Если $\text{proj}_W \vec{v} = \vec{0}$, то $\vec{v} \in W^\perp$, и угол равен $\pi/2$ (90 градусов).

6.2.4. Применение Проекций в Обучении с Подкреплением

1. **Метод наименьших квадратов (Least Squares):** Многие алгоритмы аппроксимации функций ценности в RL, такие как LSTD (Least Squares Temporal Difference) и LSPI (Least Squares Policy Iteration), основаны на решении задач наименьших квадратов. Нахождение решения задачи $A\vec{x} \approx \vec{b}$ методом наименьших квадратов эквивалентно ортогональному проектированию вектора \vec{b} на пространство столбцов матрицы A .
2. **Ортогонализация признаков:** В линейных методах аппроксимации признаков представление состояний или пар состояние-действие может быть предварительно ортогонализировано (например, с помощью процесса Грама-Шмидта, который тесно связан с проекциями). Это может улучшить обусловленность задачи и ускорить сходимость.
3. **Разложение функции ценности:** Функцию ценности можно разложить на компоненты, лежащие в различных подпространствах признаков. Например, в задачах с обобщенной аппроксимацией функции ценности (GVF - General Value Functions) проекции могут использоваться для анализа и построения этих компонент.
4. **Понижение размерности и выделение признаков:** Проекция данных на подпространство меньшей размерности (например, в методе главных компонент, PCA, который ищет подпространство, на которое проекция данных имеет максимальную дисперсию) является ключевым шагом во многих техниках понижения размерности. Это актуально для RL в пространствах состояний высокой размерности.
5. **Ограничения на стратегию или функцию ценности:** Иногда требуется, чтобы стратегия или функция ценности удовлетворяли определенным линейным ограничениям. Это можно интерпретировать как поиск решения в некотором подпространстве, и проекции могут использоваться для обеспечения выполнения этих ограничений.
6. **Residual Algorithms (Остаточные алгоритмы):** Некоторые алгоритмы в RL (например, остаточные градиентные алгоритмы) работают с "остатками" или ошибками предсказаний. Ортогональные проекции могут играть роль в анализе и минимизации этих остатков.

Понимание проекций и ортогональных дополнений дает инструменты для решения задач аппроксимации, оптимизации и анализа данных, которые часто возникают в контексте обучения с подкреплением.

6.3. Изменение Базиса и Матрица Перехода

6.3.1. Координаты Вектора в Базисе

Пусть V – векторное пространство размерности n , и пусть $B = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ – упорядоченный базис в V . Любой вектор $\vec{v} \in V$ может быть единственным образом представлен как линейная комбинация базисных векторов:

$$\vec{v} = x_1\vec{e}_1 + x_2\vec{e}_2 + \dots + x_n\vec{e}_n = \sum_{i=1}^n x_i\vec{e}_i$$

Числа (x_1, x_2, \dots, x_n) называются **координатами** вектора \vec{v} в базисе B . Вектор-столбец этих координат обозначается как $[\vec{v}]_B$:

$$[\vec{v}]_B = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

6.3.2. Матрица Перехода от Одного Базиса к Другому

Пусть в пространстве V заданы два базиса:

- "Старый" базис: $B = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$
- "Новый" базис: $B' = \{\vec{e}'_1, \vec{e}'_2, \dots, \vec{e}'_n\}$

Каждый вектор нового базиса B' можно выразить как линейную комбинацию векторов старого базиса B :

$$\vec{e}'_1 = p_{11}\vec{e}_1 + p_{21}\vec{e}_2 + \dots + p_{n1}\vec{e}_n = \sum_{i=1}^n p_{i1}\vec{e}_i$$

$$\vec{e}'_2 = p_{12}\vec{e}_1 + p_{22}\vec{e}_2 + \dots + p_{n2}\vec{e}_n = \sum_{i=1}^n p_{i2}\vec{e}_i$$

$$\vdots$$

$$\vec{e}'_n = p_{1n}\vec{e}_1 + p_{2n}\vec{e}_2 + \dots + p_{nn}\vec{e}_n = \sum_{i=1}^n p_{in}\vec{e}_i$$

Матрицей перехода от старого базиса B к новому базису B' называется матрица $P_{B \rightarrow B'}$, столбцами которой являются координатные векторы новых базисных векторов \vec{e}'_j в старом базисе B :

$$P_{B \rightarrow B'} = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix}$$

То есть, j -й столбец матрицы $P_{B \rightarrow B'}$ – это $[\vec{e}'_j]_B$.

Матрица перехода $P_{B \rightarrow B'}$ всегда является невырожденной (обратимой), так как базисные векторы линейно независимы.

6.3.3. Преобразование Координат Вектора при Замене Базиса

Пусть вектор \vec{v} имеет координаты $[\vec{v}]_B = (x_1, \dots, x_n)^T$ в старом базисе B и координаты $[\vec{v}]_{B'} = (x'_1, \dots, x'_n)^T$ в новом базисе B' .

Связь между этими координатами выражается формулой:

$$[\vec{v}]_B = P_{B \rightarrow B'} [\vec{v}]_{B'}$$

Или, для нахождения новых координат через старые:

$$[\vec{v}]_{B'} = (P_{B \rightarrow B'})^{-1} [\vec{v}]_B$$

Обозначим $P = P_{B \rightarrow B'}$. Тогда $[\vec{v}]_B = P [\vec{v}]_{B'}$ и $[\vec{v}]_{B'} = P^{-1} [\vec{v}]_B$.

Матрица $P_{B' \rightarrow B} = (P_{B \rightarrow B'})^{-1}$ является матрицей перехода от базиса B' к базису B .

Замечание по обозначениям: Иногда матрицу перехода от B к B' определяют как матрицу, j -й столбец которой содержит координаты *старого* базисного вектора \vec{e}_j в *новом* базисе B' . В этом случае формулы преобразования координат будут выглядеть иначе. Важно всегда уточнять определение матрицы перехода.

В данном изложении мы придерживаемся определения, где столбцы $P_{B \rightarrow B'}$ – это новые векторы в старом базисе.

6.3.4. Преобразование Матрицы Линейного Оператора при Замене Базиса

Пусть $A : V \rightarrow V$ – линейный оператор, и пусть M_B – матрица этого оператора в базисе B , а $M_{B'}$ – матрица этого же оператора в базисе B' .

Если $P = P_{B \rightarrow B'}$ – матрица перехода от базиса B к базису B' , то связь между матрицами оператора в этих базисах следующая:

$$M_{B'} = P^{-1} M_B P$$

Эта формула показывает, как преобразуется матрица линейного оператора при переходе к новому базису. Матрицы M_B и $M_{B'}$ называются **подобными**.

Если оператор $A : V \rightarrow W$, и в V происходит замена базиса $B_V \rightarrow B'_V$ с матрицей перехода P_V , а в W – замена базиса $B_W \rightarrow B'_W$ с матрицей перехода P_W , то матрица

оператора M_{B_V, B_W} (действующая на координаты в базисах B_V, B_W) и $M_{B'_V, B'_W}$ связаны так:

$$M_{B'_V, B'_W} = P_W^{-1} M_{B_V, B_W} P_V$$

6.3.5. Применение Изменения Базиса в Обучении с Подкреплением

- 1. Выбор удобного представления состояний/действий:** В RL пространство состояний или признаков может быть представлено в различных базисах. Переход к новому базису может упростить структуру задачи, сделать признаки более интерпретируемыми или улучшить обусловленность матриц, используемых в алгоритмах.
Например, PCA (метод главных компонент) по сути является переходом к новому ортонормированному базису (базису из собственных векторов ковариационной матрицы), в котором данные становятся некоррелированными, и можно выбрать наиболее важные компоненты.
- 2. Диагонализация матриц операторов:** Как показано выше, переход к базису из собственных векторов (если он существует) диагонализует матрицу линейного оператора. Это упрощает анализ динамики системы или свойств оператора (например, оператора Беллмана).
- 3. Упрощение аппроксимации функций:** При аппроксимации функций ценности или стратегий выбор подходящего базиса для функций-аппроксиматоров (например, полиномиальный базис, радиальные базисные функции, вейвлеты) критичен. Переход между различными базисами может быть необходим для адаптации к структуре задачи.
- 4. Анализ инвариантности:** Некоторые свойства алгоритмов RL или оптимальных стратегий могут быть инвариантны относительно определенных преобразований базиса. Понимание того, как смена базиса влияет на ключевые величины (например, Q-функции, функции преимущества), помогает в теоретическом анализе.
- 5. Декомпозиция задач:** Иногда задачу RL можно разложить на более простые подзадачи путем перехода в базис, который разделяет различные аспекты или факторы пространства состояний.
- 6. Передача знаний (Transfer Learning):** При переносе знаний из одной задачи RL в другую может потребоваться преобразование представлений состояний или действий, что можно рассматривать как изменение базиса, чтобы согласовать пространства признаков.

Понимание механики смены базиса и того, как при этом преобразуются координаты векторов и матрицы операторов, является фундаментальным для многих продвинутых методов в линейной алгебре и их приложений в RL, особенно когда речь идет об аппроксимации, анализе и упрощении сложных моделей.

6.4. Разложения Матриц: LU и QR

Разложения (или факторизации) матриц представляют собой способы представления данной матрицы в виде произведения двух или более матриц с более простой структурой (например, треугольных, диагональных, ортогональных). Эти разложения являются ключевыми инструментами в численной линейной алгебре для решения систем линейных уравнений, нахождения собственных значений, решения задач наименьших квадратов и т.д.

6.4.1. LU-разложение (LU-декомпозиция)

Определение:

LU-разложение квадратной матрицы A – это ее представление в виде произведения нижней треугольной матрицы L (Lower) и верхней треугольной матрицы U (Upper):

$$A = LU$$

- Матрица L – нижняя треугольная (все элементы выше главной диагонали равны нулю). Часто на ее главной диагонали стоят единицы (такое разложение называется разложением Дулиттла, или L_1U -разложением). Если на диагонали U стоят единицы, это разложение Краута (LU_1).
- Матрица U – верхняя треугольная (все элементы ниже главной диагонали равны нулю).

Существование и единственность:

- Если все главные миноры матрицы A (определители левых верхних подматриц) ненулевые, то LU-разложение существует и единственно (при условии, что диагональные элементы L или U фиксированы, например, $l_{ii} = 1$).
- Если какой-то главный минор равен нулю, "чистое" LU-разложение может не существовать или быть не единственным. В таких случаях используется **LU-разложение с перестановками (LUP-разложение)**:

$$PA = LU$$

где P – матрица перестановок (получается из единичной матрицы перестановкой строк). Это разложение существует для любой невырожденной матрицы A и является численно более устойчивым.

Методы нахождения LU-разложения:

Основной метод – **метод Гаусса**. Процесс приведения матрицы A к верхнетреугольному виду U с помощью элементарных преобразований строк (вычитание одной строки, умноженной на коэффициент, из другой) можно представить как умножение A слева на последовательность нижнетреугольных матриц. Матрица L тогда будет произведением обратных к этим элементарным матрицам.

Применения LU-разложения:

1. Решение систем линейных уравнений (СЛУ):

Если $A\vec{x} = \vec{b}$, то $LU\vec{x} = \vec{b}$. Решение разбивается на два этапа:

- Решить $L\vec{y} = \vec{b}$ относительно \vec{y} (прямая подстановка, так как L нижнетреугольная).
- Решить $U\vec{x} = \vec{y}$ относительно \vec{x} (обратная подстановка, так как U верхнетреугольная).

Это особенно эффективно, если нужно решить СЛУ с одной и той же матрицей A для многих различных векторов \vec{b} , так как разложение $A = LU$ выполняется один раз.

2. **Вычисление определителя матрицы:** $\det(A) = \det(L) \det(U)$. Так как L и U треугольные, их определители равны произведению диагональных элементов. Если $l_{ii} = 1$, то $\det(L) = 1$, и $\det(A) = \det(U) = \prod u_{ii}$. Для LUP-разложения: $\det(A) = (-1)^s \prod l_{ii} \prod u_{ii}$, где s – число перестановок в P .

3. **Вычисление обратной матрицы:** $A^{-1} = (LU)^{-1} = U^{-1}L^{-1}$. Обращение треугольных матриц проще, чем обращение произвольной матрицы.

6.4.2. QR-разложение (QR-декомпозиция)

Определение:

QR-разложение матрицы A (размера $m \times n$) – это ее представление в виде произведения ортогональной (или унитарной в комплексном случае) матрицы Q и верхней треугольной (или верхнетрапецевидной) матрицы R :

$$A = QR$$

- Матрица Q (размера $m \times m$) – ортогональная, т.е. $Q^T Q = Q Q^T = I_m$ (ее столбцы образуют ортонормированный базис \mathbb{R}^m). (Если рассматривается "тонкое" или "экономичное" QR-разложение, то Q будет $m \times n$ и $Q^T Q = I_n$).
- Матрица R (размера $m \times n$) является верхней треугольной (или верхнетрапецевидной, если $m > n$, то есть $R_{ij} = 0$ при $i > j$).

Существование и способы нахождения:

QR-разложение существует для любой вещественной или комплексной матрицы. Основные методы построения QR-разложения:

- Процесс ортогонализации Грама-Шмидта:** Столбцы матрицы A рассматриваются как набор векторов. Применяя к ним процесс Грама-Шмидта, мы получаем ортонормированный набор векторов, которые становятся столбцами матрицы Q . Матрица R содержит коэффициенты, выражающие исходные столбцы A через ортонормированные столбцы Q . $A_k = Q_k R_k$, где A_k и Q_k – матрицы из первых k столбцов A и Q соответственно, R_k – верхняя треугольная $k \times k$ матрица. $R_{ij} = (\vec{q}_i, \vec{a}_j)$ для $i < j$, $R_{ii} = \|\vec{a}_i - \sum_{k=1}^{i-1} R_{ki} \vec{q}_k\|$, и $R_{ij} = 0$ для $i > j$. $\vec{q}_i = (\vec{a}_i - \sum_{k=1}^{i-1} R_{ki} \vec{q}_k) / R_{ii}$.
Классический процесс Грама-Шмидта может быть численно неустойчив. Модифицированный процесс Грама-Шмидта обычно более устойчив.

2. **Преобразования Хаусхолдера (отражения Хаусхолдера):** Матрица A последовательно умножается слева на ортогональные матрицы Хаусхолдера H_1, H_2, \dots, H_n (или $\min(m-1, n)$) так, чтобы занулить поддиагональные элементы в каждом столбце, приводя A к верхнетреугольному виду R . Тогда $H_n \dots H_2 H_1 A = R$. Так как матрицы Хаусхолдера ортогональны и симметричны ($H_i^T = H_i$, $H_i^2 = I$), то $Q = H_1 H_2 \dots H_n$. Этот метод численно более устойчив, чем Грама-Шмидта.
3. **Вращения Гивенса:** Матрица A также приводится к верхнетреугольному виду R путем последовательного умножения слева на ортогональные матрицы вращения Гивенса. Каждое вращение зануляет один конкретный поддиагональный элемент. Q формируется как произведение транспонированных матриц вращения. Этот метод полезен, когда нужно занулить лишь несколько элементов или когда матрица A разреженная.

Применения QR-разложения:

1. **Решение систем линейных уравнений (особенно методом наименьших квадратов):**
 Для решения СЛУ $A\vec{x} = \vec{b}$, подставляем $A = QR$: $QR\vec{x} = \vec{b}$. Умножая слева на Q^T (так как $Q^T Q = I$), получаем $R\vec{x} = Q^T \vec{b}$. Эта система с верхнетреугольной матрицей R легко решается обратной подстановкой.
 Это особенно полезно для решения переопределенных систем ($m > n$) методом наименьших квадратов. Задача минимизации $\|A\vec{x} - \vec{b}\|^2$ сводится к решению $R\vec{x} = Q_1^T \vec{b}$, где Q_1 – первые n столбцов Q (если используется "тонкое" QR-разложение $A = Q_1 R_1$).
2. **Нахождение собственных значений и векторов (QR-алгоритм):** QR-алгоритм является одним из основных методов для вычисления собственных значений матриц. Он итеративно строит последовательность матриц $A_k = Q_k R_k$, $A_{k+1} = R_k Q_k$. При определенных условиях A_k сходится к верхнетреугольной или квазитреугольной матрице (форма Шура), диагональные элементы которой являются собственными значениями исходной матрицы A .
3. **Сингулярное разложение (SVD):** QR-разложение может использоваться как шаг в некоторых алгоритмах вычисления SVD.
4. **Ортогонализация базисов:** Сам процесс построения Q из A является способом получения ортонормированного базиса (столбцы Q) для пространства, натянутого на столбцы A .

6.4.3. Применение Разложений Матриц в Обучении с Подкреплением

1. **Решение систем уравнений Беллмана:** Многие алгоритмы RL, особенно те, что основаны на методах временных разностей (TD) или аппроксимации функции ценности, сводятся к решению больших систем линейных уравнений (например,

уравнения Беллмана в матричной форме $(I - \gamma P_\pi)V_\pi = R_\pi$. LU- и QR-разложения являются стандартными численными методами для решения таких систем.

- LSTD (Least Squares Temporal Difference) и его варианты часто используют QR-разложение или SVD для решения задачи наименьших квадратов.

2. **Аппроксимация функции ценности:** При использовании линейной архитектуры для аппроксимации функции ценности $V(s) \approx \vec{\phi}(s)^T \vec{w}$, параметры \vec{w} часто находятся решением задачи наименьших квадратов, где матричные разложения играют ключевую роль.
3. **Алгоритмы градиентов политики с натуральным градиентом:** Натуральный градиент часто включает обращение матрицы информации Фишера. Разложения матриц могут использоваться для эффективного вычисления или аппроксимации этого обращения.
4. **Понижение размерности и отбор признаков:** Хотя SVD и PCA более прямо связаны с понижением размерности, QR-разложение (особенно с выбором ведущих столбцов, QR with column pivoting) может использоваться для отбора подмножества линейно независимых или наиболее "важных" признаков.
5. **Численная стабильность:** В сложных алгоритмах RL, где производятся многочисленные матричные операции, использование численно устойчивых разложений (таких как QR на основе Хаусхолдера или Гивенса, или SVD) критически важно для предотвращения накопления ошибок и получения надежных результатов.
6. **Методы Монте-Карло для матриц:** Некоторые продвинутые методы могут использовать матричные разложения в контексте стохастических аппроксимаций или оценок, связанных с большими матрицами, возникающими в RL (например, матрицы переходных вероятностей или ковариационные матрицы).

Понимание LU- и QR-разложений, их свойств и методов вычисления дает разработчику алгоритмов RL мощные инструменты для решения численных задач, возникающих при моделировании и оптимизации в сложных средах.