Семинар 13

Выделение базиса из системы векторов

Дано Пусть $v_1, \ldots, v_m \in \mathbb{R}^n$ – вектора.

Задача Найти базис системы векторов $v_1, \ldots, v_m$ и разложить оставшиеся вектора по этому базису.

Алгоритм

1. Запишем вектора $v_1, \ldots, v_m$ по столбцам в матрицу $A \in \mathrm{M}_{n\,m}(\mathbb{R})$. Например, при $n = 3$, $m = 5$
$$A = \begin{pmatrix} v_{11} & v_{21} & v_{31} & v_{41} & v_{51} \\ v_{12} & v_{22} & v_{32} & v_{42} & v_{52} \\ v_{13} & v_{23} & v_{33} & v_{43} & v_{53} \end{pmatrix}$$

2. Приведем матрицу A элементарными преобразованиями строк к улучшенному ступенчатому виду. Например
$$A' = \begin{pmatrix} 1 & 0 & a_{31} & 0 & a_{51} \\ 0 & 1 & a_{32} & 0 & a_{52} \\ 0 & 0 & 0 & 1 & a_{53} \end{pmatrix}$$

3. Пусть $k_1, \ldots, k_r$ – номера главных позиций в матрице A'. Тогда вектора $v_{k_1}, \ldots, v_{k_r}$ образуют базис системы $v_1, \ldots, v_k$. Например, в примере выше это вектора v_1, v_2 и v_4.

4. Пусть v_i – вектор соответствует неглавной позиции в A'. Тогда в i-ом столбце A' записаны координаты разложения v_i через найденный базис выше. Например, в примере выше $v_3 = a_{31}v_1 + a_{32}v_2$ и $v_5 = a_{51}v_1 + a_{52}v_2 + a_{53}v_4$.

Корректность алгоритма

Существует два способа объяснять почему этот алгоритм корректен. Первый я показывал в предыдущий раз, теперь давайте покажу идейный метод.

Через замену координат Теперь посмотрим на эту ситуацию по-другому. Пусть у нас есть вектор $v \in F^m$. Если $v^t = (v_1, \ldots, v_m)$, то на v_i можно смотреть как на координаты v в стандартном базисе $e_1, \ldots, e_m$. Если у нас естьи невырожденная матрица $C \in \mathrm{M}_m(F)$, то мы можем сделать замену базиса по правилу $(e_1, \ldots, e_m)C^{-1} = (f_1, \ldots, f_m)$. Тогда координаты вектора v изменятся по правилу $v \mapsto Cv$. То есть в базисе $f_1, \ldots, f_m$ вектор Cv будет вектором координат. То есть v и Cv – это координаты одного и того же вектора, только в разных базисах.

Теперь посмотрим на систему векторов $v_1, \ldots, v_n \in F^m$. Если я с матрицей $A = (v_1|\ldots|v_n)$ сделаю элементарные преобразования строк, это равносильно умножению слева на невырожденную матрицу C. То есть матрица A заменится на $A' = (Cv_1|\ldots,|Cv_n)$ (опять по блочным формулам). Но тогда на столбцы матрицы A' можно смотреть как на те же вектора, что и в столбцах матрицы A, только записанные в другом базисе. А раз это те же векторы, что и были, но в более удобных координатах, то решать задачу для столбцов A – это то же самое, что решать задачу для столбцов A'. Тут надо обратить внимание, что понятие линейной зависимости или независимости никак не связано с выбором координат для векторов.

Ранг системы векторов

Пусть V – некоторое векторное пространство. Системой векторов называется последовательность $(v_1, \ldots, v_k)$ из векторов V, в которой векторы v_i могут повторяться.[1]

По определению рангом системы $(v_1, \ldots, v_k)$ называется максимальное количество линейно независимых векторов в этой системе. Ранг такой системы будет обозначаться $\mathrm{rk}(v_1, \ldots, v_k)$.

Утверждение. *Если $(v_1, \ldots, v_k)$ – некоторая система векторов в векторном пространстве V, то* $\mathrm{rk}(v_1, \ldots, v_k) = \dim\langle v_1, \ldots, v_k\rangle$.

[1] В подобной ситуации повторяющиеся векторы различаются по индексу – «ключу».

Матричный ранг

Пусть $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ – некоторая матрица. Сейчас я определю пять разных определений ранга матрицы. Все эти ранги между собой совпадают и полученная величина будет просто называться рангом матрицы A и обозначаться $\operatorname{rk} A$.

Определение. Пусть $A_1, \ldots, A_n \in \mathbb{R}^m$ – столбцы матрицы A, то есть $A = (A_1|\ldots|A_n)$. Тогда столбцовым рангом матрицы A называется ранг системы $(A_1, \ldots, A_n)$, то есть $\operatorname{rk}_{\text{столб}} A = \operatorname{rk}(A_1, \ldots, A_n)$.

Определение. Пусть $A_1, \ldots, A_m \in \mathbb{R}^n$ – строки матрицы A, то есть $A^t = (A_1|\ldots|A_m)$. Тогда строковым рангом матрицы A называется ранг системы $(A_1, \ldots, A_m)$, то есть $\operatorname{rk}_{\text{стр}} A = \operatorname{rk}(A_1, \ldots, A_m)$.

Определение. Факториальным рангом матрицы A называется следующее число

$$\min\{k \mid A = BC, \text{ где } B \in \mathrm{M}_{m\,k}(\mathbb{R}),\, C \in \mathrm{M}_{k\,n}(\mathbb{R})\}$$

то есть это минимальное число k такое, что матрица A представима в виде произведения матриц BC, где общая размерность для B и C, по которой они перемножаются, есть k.

Определение. Тензорным рангом матрицы A называется следующее число

$$\min\{k \mid A = x_1y_1^t + \ldots + x_ky_k^t, \text{ где } x_i \in \mathbb{R}^m,\, y_i \in \mathbb{R}^n\}$$

то есть это минимальное число k такое, что матрица A представима в виде суммы k «тощих» матриц вида xy^t, где $x \in \mathbb{R}^m$ и $y \in \mathbb{R}^n$.

Если я в матрице A выделю какой-нибудь набор из k строк и одновременно набор из k столбцов, а потом возьму матрицу составленную из элементов на пересечении этих строк и столбцов, то я получу квадратную матрицу размера k. Такие матрицы мы будем называть квадратными подматрицами матрицы A.

Определение. Минорным рангом матрицы A называется размер наибольшей невырожденной квадратной подматрицы.[2]

Главное для нас следующее утверждение.

Утверждение. *Для любой матрицы $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ все пять видов ранга совпадают и не превосходят* $\min(m, n)$.

Примеры

1. В начале заметим, что матрица имеет ранг 0 тогда и только тогда, когда $A = 0$.

2. Ранг матрицы A равен единице тогда и только тогда, когда она не нулевая и все столбцы пропорциональны одному общему столбцу (или что эквивалентно, все строки пропорциональны одной общей строке). Если воспользоваться определением факториального ранга, то мы видим, что тогда матрица A имеет вид $A = xy^t$, где $x \in \mathbb{R}^m$ и $y \in \mathbb{R}^n$ – ненулевые вектора.

Свойства ранга

Прежде всего надо запомнить как ранг связан с матричными операциями.

Утверждение. *Пусть $A, B \in \mathrm{M}_{m\,n}(\mathbb{R})$, тогда*

$$|\operatorname{rk} A - \operatorname{rk} B| \leqslant \operatorname{rk}(A + B) \leqslant \operatorname{rk} A + \operatorname{rk} B$$

Надо понимать, что, во-первых, все эти эффекты можно увидеть на диагональных матрицах; во-вторых, все границы неравенств достигаются. Смысл этого утверждения вот в чем: если вы шевелите матрицу A с помощью матрицы B, то ранг A может измениться не более чем на ранг B в любую сторону. Теперь посмотрим на матрицы: $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ и $C = -A$. Тогда $\operatorname{rk}(A + B) = \operatorname{rk} A + \operatorname{rk} B$ и $\operatorname{rk}(A + C) = \operatorname{rk} A - \operatorname{rk} C$.

[2] На самом деле можно дать более сильное определение, а именно, минорный ранг – это размер любой максимальной невырожденной подматрицы. То есть мы берем какую-то квадратную подматрицу, которая невырождена, а любая большая подматрица уже вырождена. Оказывается, что все максимальные невырожденные подматрицы имеют одинаковый размер и он называется минорным рангом.

Утверждение. *Пусть* $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ *и* $B \in \mathrm{M}_{n\,k}(\mathbb{R})$*, тогда*

$$\operatorname{rk} A + \operatorname{rk} B - n \leqslant \operatorname{rk}(AB) \leqslant \min(\operatorname{rk} A, \operatorname{rk} B)$$

Как и в предыдущем случае, все обе границы неравенства достигаются и все можно пронаблюдать на диагональных матрицах. Пусть $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ и $B = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Тогда $\operatorname{rk}(AA) = \operatorname{rk} A$ и $\operatorname{rk}(AB) = \operatorname{rk} A + \operatorname{rk} B - 2$.

Утверждение. *Пусть* $A \in \mathrm{M}_n(\mathbb{R})$ *– квадратная матрица. Тогда* $\operatorname{rk} A = n$ *тогда и только тогда, когда* A *невырождена, т.е.* $\det A \neq 0$.

Таким образом на ранг можно смотреть как на степень невырожденности матрицы A. Самый высокий ранг у невырожденных матриц, самый маленький у нулевой, но есть еще и промежуточные состояния.

Утверждение. *Если матрица* $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ *находится в ступенчатом виде и имеет* k *ступенек, то ее ранг равен* k.

Это утверждение вместе со следующим дают эффективный способ считать ранг.

Утверждение. *Для любой матрицы* $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ *и любых невырожденных матриц* $C \in \mathrm{M}_m(\mathbb{R})$ *и* $D \in \mathrm{M}_n(\mathbb{R})$ *верно:* $\operatorname{rk} A = \operatorname{rk}(CA) = \operatorname{rk}(AD)$.[3]

В частности ранг матрицы не меняется при элементарных преобразованиях столбцов и строк. Обычно этим пользуются для нахождения ранга. Более того, если $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ имеет ранг r, то элементарными преобразованиями строк и столбцов она приводится к виду

$$A \mapsto \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix}, \text{ где } E \in \mathrm{M}_r(\mathbb{R}) \text{ – единичная матрица}$$

Следствием данного замечания является следующее.

Утверждение. *Для любых матриц* $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ *и* $B \in \mathrm{M}_{s\,t}(\mathbb{R})$ *имеем*

$$\operatorname{rk} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \operatorname{rk} A + \operatorname{rk} B$$

Факториальный ранг

Пусть задана матрица $A \in \mathrm{M}_{m\,n}(\mathbb{R})$. Мы знаем, что если ее ранг $\operatorname{rk} A = r$, то найдется разложение вида $A = BC$, где $B\,\mathrm{M}_{m\,r}(\mathbb{R})$ и $C \in \mathrm{M}_{r\,n}(\mathbb{R})$. Нам бы хотелось научиться решать задачу поиска подобного разложения. Так же хотелось бы описать все такие разложения каким-нибудь образом. Давайте сразу отметим, что если мы нарежем матрицу B на столбцы, то есть $B = (B_1|\ldots|B_r)$, а матрицу C на строки $C = \begin{pmatrix} C_1 \\ \vdots \\ C_r \end{pmatrix}$, то мы сразу же получаем разложение, на котором достигается тензорный ранг $A = B_1C_1 + \ldots + B_rC_r$. Потому достаточно решать задачу для факториального ранга.

Описание разложений Давайте покажем, что

Утверждение. *Для матрицы* $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ *на ее разложении* $A = BC$*, где* $B\,\mathrm{M}_{m\,r}(\mathbb{R})$ *и* $C \in \mathrm{M}_{r\,n}(\mathbb{R})$*, достигается факториальный ранг тогда и только тогда, когда* $\operatorname{rk} B = \operatorname{rk} C = r$.

То есть все столбцы матрицы B должны быть линейно независимы и все строки матрицы C должны быть линейно независимы и оказывается, что это критерий.

Доказательство. ($\Rightarrow$) Это доказывается через оценку на ранг, что была на лекции, а именно $\operatorname{rk}(BC) \leqslant \min(\operatorname{rk} B, \operatorname{rk} C)$.

($\Leftarrow$) Раз такое разложение есть, то $\operatorname{rk} A \leqslant r$ по определению факториального ранга. Теперь нам лишь надо оценить ранг снизу через r. Для этого нам достаточно найти r линейно независимых столбцов матрицы A. Для этого посмотрим на матрицу C. Ее ранг равен r, тогда по минорному рангу найдется невырожденная подматрица размера r. Пусть это подматрица стоящая на столбцах с номерами $i_1, \ldots, i_r$. Тогда рассмотрим в C подматрицу состоящую из этих столбцов и обозначим ее $\bar{C}$. Тогда получим $(A_{i_1}|\ldots|A_{i_r}) = B\bar{C}$. И при этом $\bar{C}$ обратима. А значит $B = (A_{i_1}|\ldots|A_{i_r})\bar{C}^{-1}$. Столбцы матрицы B линейно независимы и они выражаются через столбцы с номерами $i_1, \ldots, i_k$ в матрице A. А это и означает, что столбцы под номерами $i_1, \ldots, i_k$ в матрице A линейно независимы. □

[3] В частности ранг не меняется при элементарных преобразованиях строк и столбцов.

Скелетное разложение

Один из способов найти разложение на котором достигается тензорный ранг доставляет так называемое скелетное разложение. В этом случае матрица B будет состоять целиком из базисных столбцов матрицы A.

Дано Матрица $A \in \mathrm{M}_{m\,n}(\mathbb{R})$.

Задача Найти $B\,\mathrm{M}_{m\,r}(\mathbb{R})$ и $C \in \mathrm{M}_{r\,n}(\mathbb{R})$ такие, что $A = BC$ это разложение на котором достигается факториальный ранг.

Алгоритм

1. Применим к столбцам матрицы A алгоритм поиска базиса и разложения остальных столбцов по этому базису. В результате этого алгоритма мы получим базисные столбцы $A_{i_1}, \ldots, A_{i_r}$ и матрицу улучшенного ступенчатого вида $C \in \mathrm{M}_{r\,n}(\mathbb{R})$.
2. Составим матрицу $B = (A_{i_1}, \ldots, A_{i_r})$. Тогда $A = BC$ и есть искомое разложение.

Давайте объясним корректность алгоритма на примере $m = 3$ и $n = 5$. Пусть нам дана матрица:

$$A = \begin{pmatrix} v_{11} & v_{21} & v_{31} & v_{41} & v_{51} \\ v_{12} & v_{22} & v_{32} & v_{42} & v_{52} \\ v_{13} & v_{23} & v_{33} & v_{43} & v_{53} \end{pmatrix}$$

И после приведения к улучшенному ступенчатому виду мы имеем:

$$C = \begin{pmatrix} 1 & 0 & a_{31} & 0 & a_{51} \\ 0 & 1 & a_{32} & 0 & a_{52} \\ 0 & 0 & 0 & 1 & a_{53} \end{pmatrix}$$

Значит базисные столбцы будут A_1, A_2, A_4. В этом случае $B = (A_1, A_2, A_4)$. Проверим, что BC совпадает с A. Перемножим

$$\begin{pmatrix} A_1 & A_2 & A_4 \end{pmatrix} \begin{pmatrix} 1 & 0 & a_{31} & 0 & a_{51} \\ 0 & 1 & a_{32} & 0 & a_{52} \\ 0 & 0 & 0 & 1 & a_{53} \end{pmatrix} = \begin{pmatrix} A_1 & A_2 & * & A_4 & * \end{pmatrix}$$

В начале мы видим, что базисные столбцы будут ровно те, что надо. Теперь посмотрим, что будет стоять на свободных позициях:

$$\begin{pmatrix} A_1 & A_2 & A_4 \end{pmatrix} \begin{pmatrix} 1 & 0 & a_{31} & 0 & a_{51} \\ 0 & 1 & a_{32} & 0 & a_{52} \\ 0 & 0 & 0 & 1 & a_{53} \end{pmatrix} = \begin{pmatrix} A_1 & A_2 & a_{31}A_1 + a_{32}A_2 & A_4 & a_{51}A_1 + a_{52}A_2 + a_{53}A_4 \end{pmatrix}$$

Но это в точности разложения столбцов A_3 и A_5 по базису A_1, A_2, A_4. То есть мы действительно получили матрицу A.