

Тема VI: Евклидовы и унитарные пространства

§3. Метод наименьших квадратов

Б.М.Верников М.В.Волков

Уральский федеральный университет
Институт естественных наук и математики
кафедра алгебры и фундаментальной информатики

2021/2022 учебный год

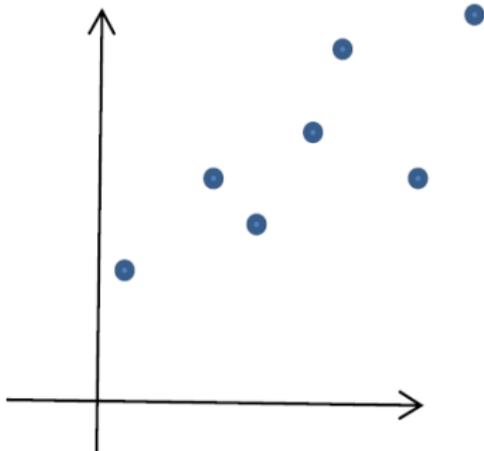
Обсудим решение несовместных систем линейных уравнений подробнее.

Обсудим решение несовместных систем линейных уравнений подробнее.

Типичный источник таких систем — обработка экспериментальных данных.

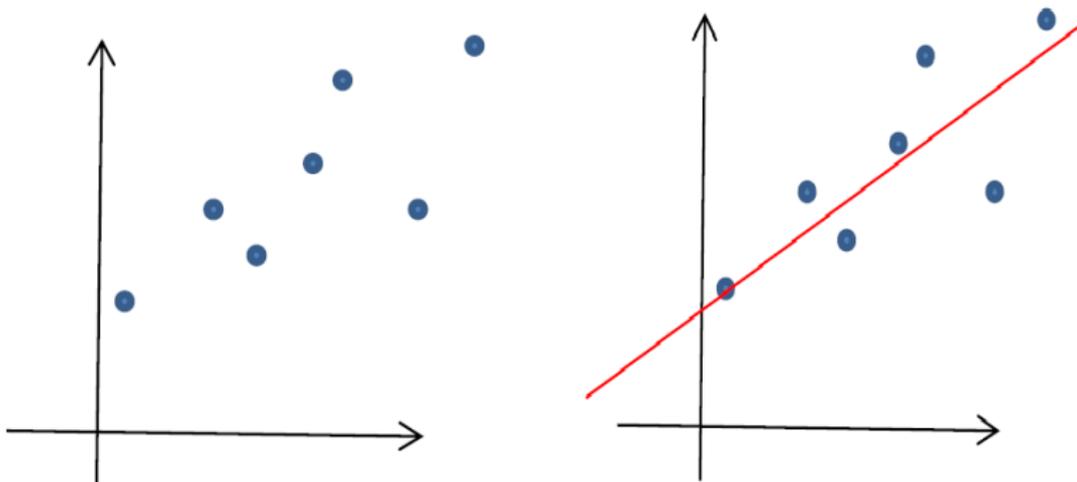
Обсудим решение несовместных систем линейных уравнений подробнее.

Типичный источник таких систем — обработка экспериментальных данных.



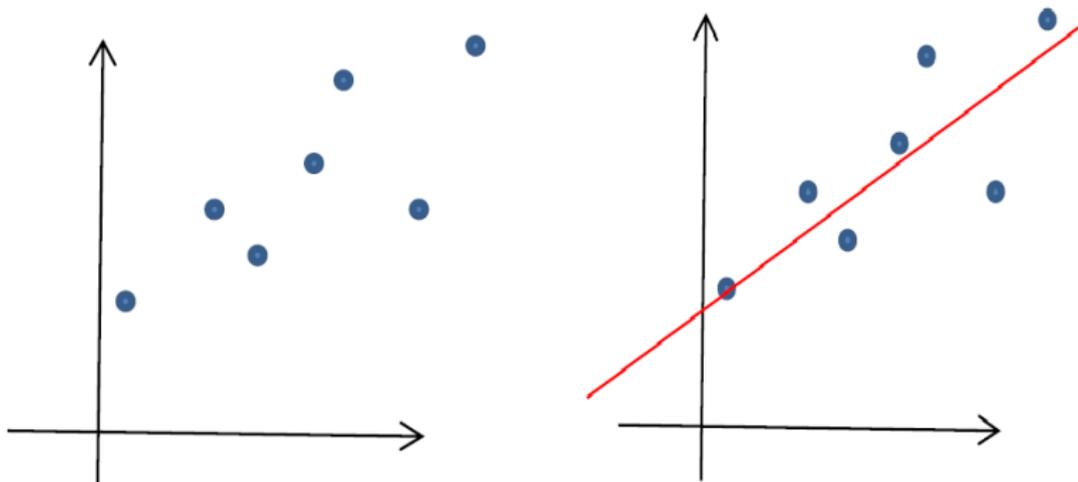
Обсудим решение несовместных систем линейных уравнений подробнее.

Типичный источник таких систем — обработка экспериментальных данных.



Обсудим решение несовместных систем линейных уравнений подробнее.

Типичный источник таких систем — обработка экспериментальных данных.



Число неизвестных мало (в данном примере ищется прямая $y = ax + b$, и неизвестные — это коэффициенты a и b), а число уравнений велико (в данном примере каждая точка (x_i, y_i) задает уравнение $y_i = ax_i + b$).

Псевдорешение системы линейных уравнений $Ax = b$ — это вектор x_0 , минимизирующий расстояние между векторами Ax и b .

Псевдорешение системы линейных уравнений $Ax = b$ — это вектор x_0 , минимизирующий расстояние между векторами Ax и b .

В конце прошлой лекции был намечен такой план поиска псевдорешений:

Псевдорешение системы линейных уравнений $Ax = \mathbf{b}$ — это вектор \mathbf{x}_0 , минимизирующий расстояние между векторами $A\mathbf{x}$ и \mathbf{b} .

В конце прошлой лекции был намечен такой план поиска псевдорешений:

- найти ортогональную проекцию \mathbf{b}_S вектора \mathbf{b} на образ S линейного отображения $x \mapsto Ax$ (т.е. на подпространство, порожденное столбцами матрицы A);

Псевдорешение системы линейных уравнений $Ax = \mathbf{b}$ — это вектор \mathbf{x}_0 , минимизирующий расстояние между векторами Ax и \mathbf{b} .

В конце прошлой лекции был намечен такой план поиска псевдорешений:

- найти ортогональную проекцию \mathbf{b}_S вектора \mathbf{b} на образ S линейного отображения $x \mapsto Ax$ (т.е. на подпространство, порожденное столбцами матрицы A);
- решить совместную систему $Ax = \mathbf{b}_S$.

Псевдорешение системы линейных уравнений $Ax = b$ — это вектор x_0 , минимизирующий расстояние между векторами Ax и b .

В конце прошлой лекции был намечен такой план поиска псевдорешений:

- найти ортогональную проекцию b_S вектора b на образ S линейного отображения $x \mapsto Ax$ (т.е. на подпространство, порожденное столбцами матрицы A);
- решить совместную систему $Ax = b_S$.

Было показано, что любое решение системы $Ax = b_S$ действительно является псевдорешением исходной системы $Ax = b$, так как наименьшее расстояние от вектора b до подпространства S есть расстояние от b до b_S .

Псевдорешение системы линейных уравнений $Ax = b$ — это вектор x_0 , минимизирующий расстояние между векторами Ax и b .

В конце прошлой лекции был намечен такой план поиска псевдорешений:

- найти ортогональную проекцию b_S вектора b на образ S линейного отображения $x \mapsto Ax$ (т.е. на подпространство, порожденное столбцами матрицы A);
- решить совместную систему $Ax = b_S$.

Было показано, что любое решение системы $Ax = b_S$ действительно является псевдорешением исходной системы $Ax = b$, так как наименьшее расстояние от вектора b до подпространства S есть расстояние от b до b_S .

В реальных задачах ранг матрицы A равен числу неизвестных.

Псевдорешение системы линейных уравнений $Ax = b$ — это вектор x_0 , минимизирующий расстояние между векторами Ax и b .

В конце прошлой лекции был намечен такой план поиска псевдорешений:

- найти ортогональную проекцию b_S вектора b на образ S линейного отображения $x \mapsto Ax$ (т.е. на подпространство, порожденное столбцами матрицы A);
- решить совместную систему $Ax = b_S$.

Было показано, что любое решение системы $Ax = b_S$ действительно является псевдорешением исходной системы $Ax = b$, так как наименьшее расстояние от вектора b до подпространства S есть расстояние от b до b_S .

В реальных задачах ранг матрицы A равен числу неизвестных. (Например, в задаче проведения прямой $y = ax + b$ через набор точек $\{(x_i, y_i)\}$ ранг равен 1 только, если все эти точки лежат на одной вертикальной прямой.)

Псевдорешение системы линейных уравнений $Ax = b$ — это вектор x_0 , минимизирующий расстояние между векторами Ax и b .

В конце прошлой лекции был намечен такой план поиска псевдорешений:

- найти ортогональную проекцию b_S вектора b на образ S линейного отображения $x \mapsto Ax$ (т.е. на подпространство, порожденное столбцами матрицы A);
- решить совместную систему $Ax = b_S$.

Было показано, что любое решение системы $Ax = b_S$ действительно является псевдорешением исходной системы $Ax = b$, так как наименьшее расстояние от вектора b до подпространства S есть расстояние от b до b_S .

В реальных задачах ранг матрицы A равен числу неизвестных.

При таком условии система $Ax = b_S$ имеет единственное решение, а значит, имеется *единственное* псевдорешение исходной системы $Ax = b$.

Псевдорешение системы линейных уравнений $Ax = b$ — это вектор x_0 , минимизирующий расстояние между векторами Ax и b .

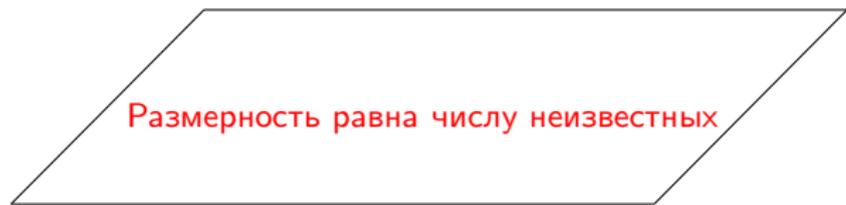
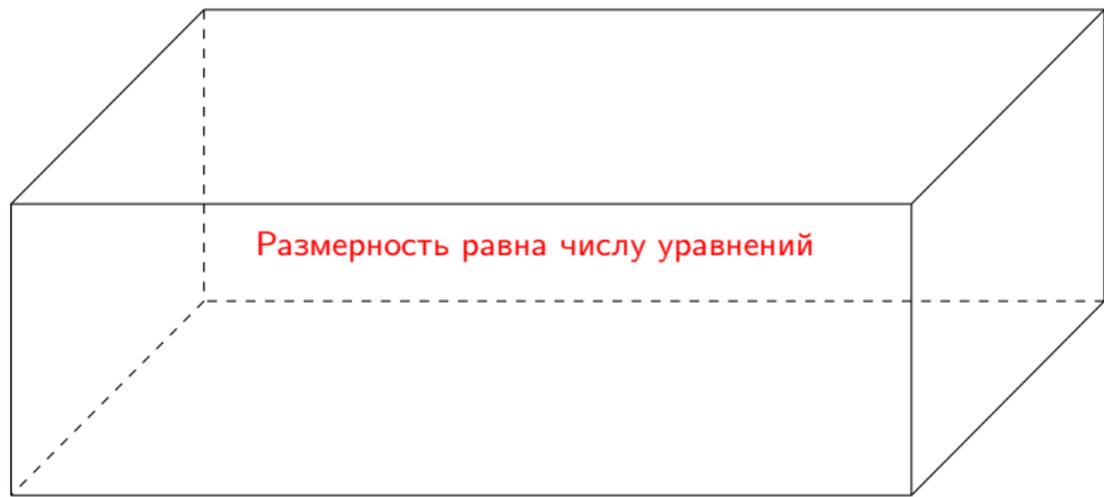
В конце прошлой лекции был намечен такой план поиска псевдорешений:

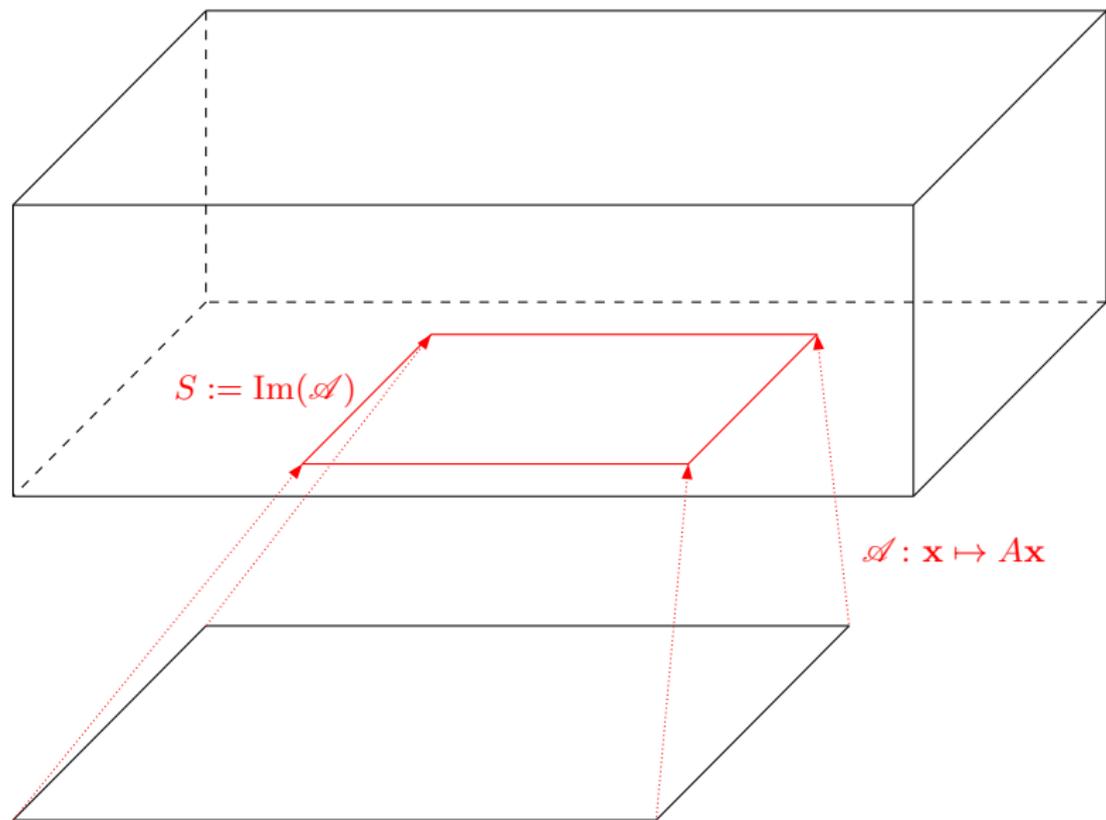
- найти ортогональную проекцию b_S вектора b на образ S линейного отображения $x \mapsto Ax$ (т.е. на подпространство, порожденное столбцами матрицы A);
- решить совместную систему $Ax = b_S$.

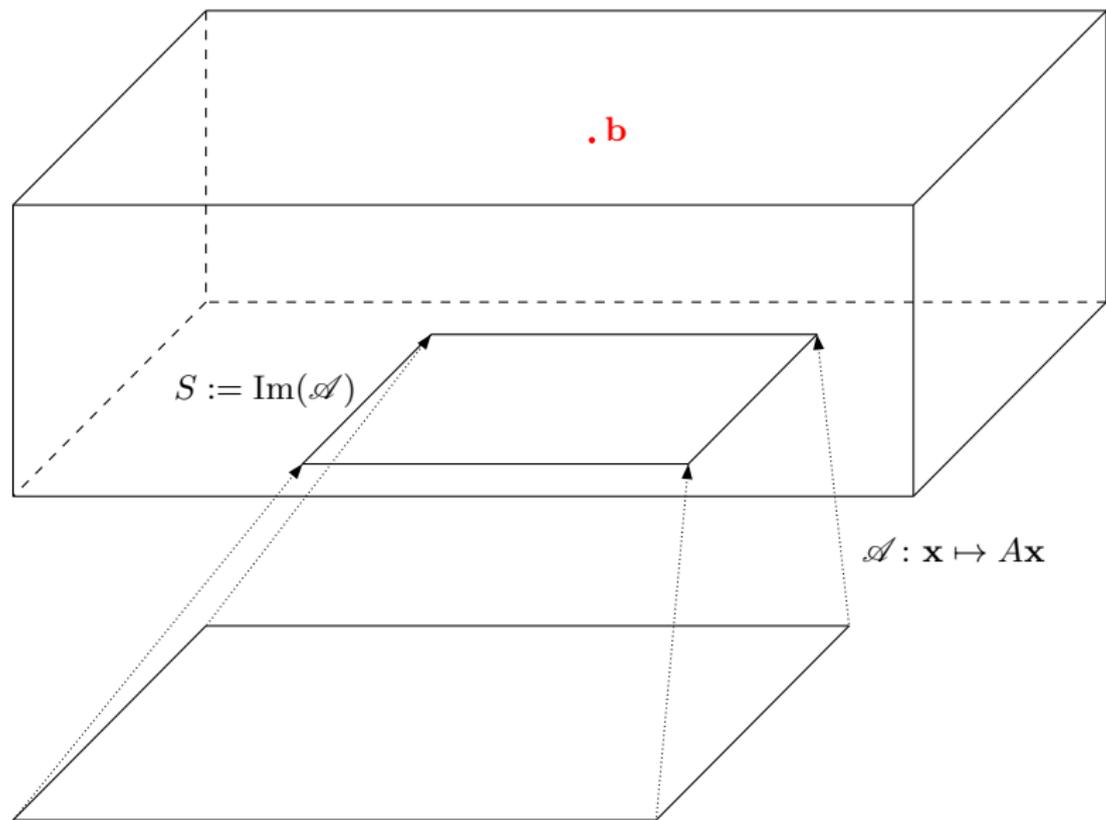
Было показано, что любое решение системы $Ax = b_S$ действительно является псевдорешением исходной системы $Ax = b$, так как наименьшее расстояние от вектора b до подпространства S есть расстояние от b до b_S .

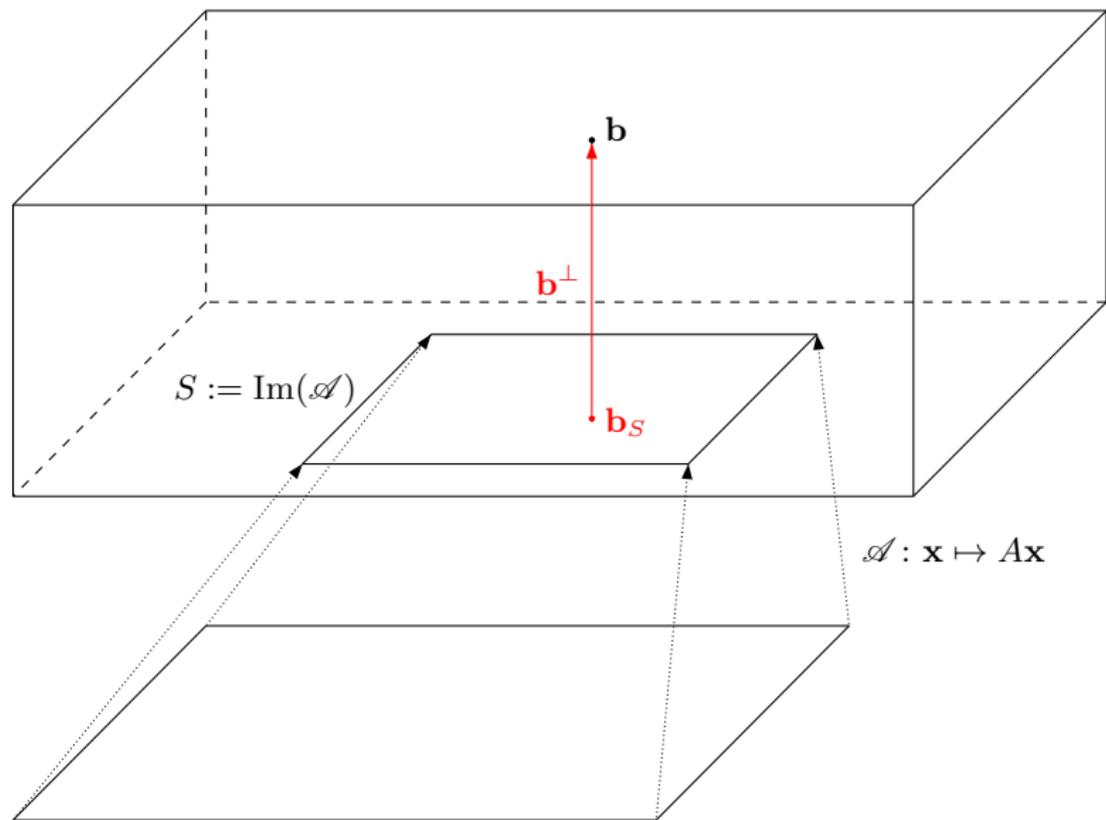
В реальных задачах ранг матрицы A равен числу неизвестных. При таком условии система $Ax = b_S$ имеет единственное решение, а значит, имеется *единственное* псевдорешение исходной системы $Ax = b$.

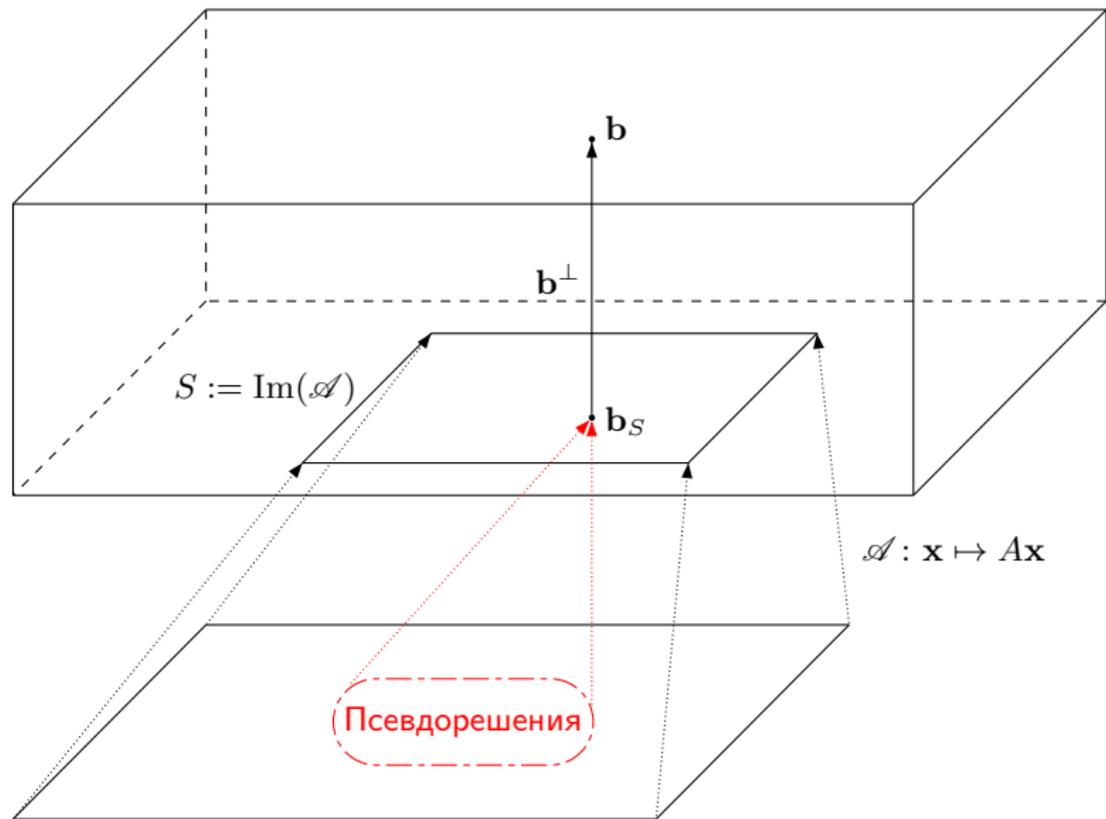
Если псевдорешение неединственно, то обычно интересуются псевдорешением наименьшей длины (*нормальное* псевдорешение). К вопросу о нормальных псевдорешениях мы вернемся позже.











Концептуально подход, описанный и проиллюстрированный выше, прост.

Концептуально подход, описанный и проиллюстрированный выше, прост. Однако вычисление ортогональной проекции с помощью процесса Грама–Шмидта приводит к громоздким и *неустойчивым* вычислениям.

Концептуально подход, описанный и проиллюстрированный выше, прост. Однако вычисление ортогональной проекции с помощью процесса Грама–Шмидта приводит к громоздким и *неустойчивым* вычислениям.

Опишем простое соображение, которое позволяет находить псевдорешения без вычисления ортогональной проекции. Его называют *методом наименьших квадратов*, поскольку речь идет о минимизации длины вектора $Ax - b$, т.е. минимизации скалярного квадрата этого вектора.

Концептуально подход, описанный и проиллюстрированный выше, прост. Однако вычисление ортогональной проекции с помощью процесса Грама–Шмидта приводит к громоздким и *неустойчивым* вычислениям.

Опишем простое соображение, которое позволяет находить псевдорешения без вычисления ортогональной проекции. Его называют *методом наименьших квадратов*, поскольку речь идет о минимизации длины вектора $Ax - b$, т.е. минимизации скалярного квадрата этого вектора.

Для определенности ограничимся случаем евклидова пространства; именно этот случай важен для практики.

Концептуально подход, описанный и проиллюстрированный выше, прост. Однако вычисление ортогональной проекции с помощью процесса Грама–Шмидта приводит к громоздким и *неустойчивым* вычислениям.

Опишем простое отображение, которое позволяет находить псевдорешения без вычисления ортогональной проекции. Его называют *методом наименьших квадратов*, поскольку речь идет о минимизации длины вектора $Ax - b$, т.е. минимизации скалярного квадрата этого вектора.

Для определенности ограничимся случаем евклидова пространства; именно этот случай важен для практики.

Теорема (обоснование метода наименьших квадратов)

Пусть A — $k \times n$ -матрица над \mathbb{R} , а S — образ линейного отображения $x \mapsto Ax$ пространства \mathbb{R}^n в пространство \mathbb{R}^k . Для произвольного вектора $b \in \mathbb{R}^k$ системы линейных уравнений $Ax = b_S$ и $A^T Ax = A^T b$ равносильны.

Концептуально подход, описанный и проиллюстрированный выше, прост. Однако вычисление ортогональной проекции с помощью процесса Грама–Шмидта приводит к громоздким и *неустойчивым* вычислениям.

Опишем простое отображение, которое позволяет находить псевдорешения без вычисления ортогональной проекции. Его называют *методом наименьших квадратов*, поскольку речь идет о минимизации длины вектора $Ax - b$, т.е. минимизации скалярного квадрата этого вектора.

Для определенности ограничимся случаем евклидова пространства; именно этот случай важен для практики.

Теорема (обоснование метода наименьших квадратов)

Пусть A — $k \times n$ -матрица над \mathbb{R} , а S — образ линейного отображения $x \mapsto Ax$ пространства \mathbb{R}^n в пространство \mathbb{R}^k . Для произвольного вектора $b \in \mathbb{R}^k$ системы линейных уравнений $Ax = b_S$ и $A^T Ax = A^T b$ равносильны.

Таким образом, вектор $x \in \mathbb{R}^n$ будет псевдорешением системы $Ax = b$ тогда и только тогда, когда x является решением системы $A^T Ax = A^T b$.

Доказательство. Подпространство S порождается образами базисных векторов пространства \mathbb{R}^n , т.е. столбцами матрицы A . Столбцы матрицы A — это строки матрицы A^T .

Доказательство. Подпространство S порождается образами базисных векторов пространства \mathbb{R}^n , т.е. столбцами матрицы A . Столбцы матрицы A — это строки матрицы A^T . Из ортогонального разложения пространства \mathbb{R}^k относительно подпространства S имеем $\mathbf{b} = \mathbf{b}_S + \mathbf{b}^\perp$.

Доказательство. Подпространство S порождается образами базисных векторов пространства \mathbb{R}^n , т.е. столбцами матрицы A . Столбцы матрицы A — это строки матрицы A^T . Из ортогонального разложения пространства \mathbb{R}^k относительно подпространства S имеем $\mathbf{b} = \mathbf{b}_S + \mathbf{b}^\perp$. Умножая это равенство слева на матрицу A^T и вспоминая формулу, выражающую скалярное произведение в евклидовом пространстве через координаты в ортонормированном базисе (т.е. формулу $\mathbf{u}\mathbf{v} = [\mathbf{u}]^T[\mathbf{v}]$), получаем:

$$A^T\mathbf{b} = A^T(\mathbf{b}_S + \mathbf{b}^\perp) = A^T\mathbf{b}_S + A^T\mathbf{b}^\perp = A^T\mathbf{b}_S, \quad (\star)$$

поскольку вектор \mathbf{b}^\perp ортогонален всем векторам из S .

Доказательство. Подпространство S порождается образами базисных векторов пространства \mathbb{R}^n , т.е. столбцами матрицы A . Столбцы матрицы A — это строки матрицы A^T . Из ортогонального разложения пространства \mathbb{R}^k относительно подпространства S имеем $\mathbf{b} = \mathbf{b}_S + \mathbf{b}^\perp$. Умножая это равенство слева на матрицу A^T и вспоминая формулу, выражающую скалярное произведение в евклидовом пространстве через координаты в ортонормированном базисе, получаем:

$$A^T \mathbf{b} = A^T (\mathbf{b}_S + \mathbf{b}^\perp) = A^T \mathbf{b}_S + A^T \mathbf{b}^\perp = A^T \mathbf{b}_S, \quad (\star)$$

поскольку вектор \mathbf{b}^\perp ортогонален всем векторам из S .
Из (\star) , если $\mathbf{x} \in \mathbb{R}^n$ — решение системы $A\mathbf{x} = \mathbf{b}_S$, то

$$A^T A\mathbf{x} = A^T \mathbf{b}_S \stackrel{(\star)}{=} A^T \mathbf{b}.$$

Доказательство. Подпространство S порождается образами базисных векторов пространства \mathbb{R}^n , т.е. столбцами матрицы A . Столбцы матрицы A — это строки матрицы A^T . Из ортогонального разложения пространства \mathbb{R}^k относительно подпространства S имеем $\mathbf{b} = \mathbf{b}_S + \mathbf{b}^\perp$. Умножая это равенство слева на матрицу A^T и вспоминая формулу, выражающую скалярное произведение в евклидовом пространстве через координаты в ортонормированном базисе, получаем:

$$A^T \mathbf{b} = A^T (\mathbf{b}_S + \mathbf{b}^\perp) = A^T \mathbf{b}_S + A^T \mathbf{b}^\perp = A^T \mathbf{b}_S, \quad (*)$$

поскольку вектор \mathbf{b}^\perp ортогонален всем векторам из S .
Из (*), если $\mathbf{x} \in \mathbb{R}^n$ — решение системы $A\mathbf{x} = \mathbf{b}_S$, то

$$A^T A\mathbf{x} = A^T \mathbf{b}_S \stackrel{(*)}{=} A^T \mathbf{b}.$$

Обратно, пусть $\mathbf{x} \in \mathbb{R}^n$ является решением системы $A^T A\mathbf{x} = A^T \mathbf{b}$.

Доказательство. Подпространство S порождается образами базисных векторов пространства \mathbb{R}^n , т.е. столбцами матрицы A . Столбцы матрицы A — это строки матрицы A^T . Из ортогонального разложения пространства \mathbb{R}^k относительно подпространства S имеем $\mathbf{b} = \mathbf{b}_S + \mathbf{b}^\perp$. Умножая это равенство слева на матрицу A^T и вспоминая формулу, выражающую скалярное произведение в евклидовом пространстве через координаты в ортонормированном базисе, получаем:

$$A^T \mathbf{b} = A^T (\mathbf{b}_S + \mathbf{b}^\perp) = A^T \mathbf{b}_S + A^T \mathbf{b}^\perp = A^T \mathbf{b}_S, \quad (*)$$

поскольку вектор \mathbf{b}^\perp ортогонален всем векторам из S .
Из (*), если $\mathbf{x} \in \mathbb{R}^n$ — решение системы $A\mathbf{x} = \mathbf{b}_S$, то

$$A^T A\mathbf{x} = A^T \mathbf{b}_S \stackrel{(*)}{=} A^T \mathbf{b}.$$

Обратно, пусть $\mathbf{x} \in \mathbb{R}^n$ является решением системы $A^T A\mathbf{x} = A^T \mathbf{b}$. Возьмем произвольный вектор $\mathbf{y} \in \mathbb{R}^n$ и перемножим вектора $A\mathbf{y}$ и $A\mathbf{x} - \mathbf{b}_S$.

Доказательство. Подпространство S порождается образами базисных векторов пространства \mathbb{R}^n , т.е. столбцами матрицы A . Столбцы матрицы A — это строки матрицы A^T . Из ортогонального разложения пространства \mathbb{R}^k относительно подпространства S имеем $\mathbf{b} = \mathbf{b}_S + \mathbf{b}^\perp$. Умножая это равенство слева на матрицу A^T и вспоминая формулу, выражающую скалярное произведение в евклидовом пространстве через координаты в ортонормированном базисе, получаем:

$$A^T \mathbf{b} = A^T (\mathbf{b}_S + \mathbf{b}^\perp) = A^T \mathbf{b}_S + A^T \mathbf{b}^\perp = A^T \mathbf{b}_S, \quad (*)$$

поскольку вектор \mathbf{b}^\perp ортогонален всем векторам из S .
Из (*), если $\mathbf{x} \in \mathbb{R}^n$ — решение системы $A\mathbf{x} = \mathbf{b}_S$, то

$$A^T A\mathbf{x} = A^T \mathbf{b}_S \stackrel{(*)}{=} A^T \mathbf{b}.$$

Обратно, пусть $\mathbf{x} \in \mathbb{R}^n$ является решением системы $A^T A\mathbf{x} = A^T \mathbf{b}$. Возьмем произвольный вектор $\mathbf{y} \in \mathbb{R}^n$ и перемножим вектора $A\mathbf{y}$ и $A\mathbf{x} - \mathbf{b}_S$. Имеем (снова используя формулу $\mathbf{u}\mathbf{v} = [\mathbf{u}]^T[\mathbf{v}]$)

$$(A\mathbf{y})^T (A\mathbf{x} - \mathbf{b}_S) = \mathbf{y}^T A^T (A\mathbf{x} - \mathbf{b}_S) = \mathbf{y}^T (A^T A\mathbf{x} - A^T \mathbf{b}_S) \stackrel{(*)}{=} \mathbf{y}^T (A^T A\mathbf{x} - A^T \mathbf{b}) = 0.$$

Доказательство. Подпространство S порождается образами базисных векторов пространства \mathbb{R}^n , т.е. столбцами матрицы A . Столбцы матрицы A — это строки матрицы A^T . Из ортогонального разложения пространства \mathbb{R}^k относительно подпространства S имеем $\mathbf{b} = \mathbf{b}_S + \mathbf{b}^\perp$. Умножая это равенство слева на матрицу A^T и вспоминая формулу, выражающую скалярное произведение в евклидовом пространстве через координаты в ортонормированном базисе, получаем:

$$A^T \mathbf{b} = A^T (\mathbf{b}_S + \mathbf{b}^\perp) = A^T \mathbf{b}_S + A^T \mathbf{b}^\perp = A^T \mathbf{b}_S, \quad (*)$$

поскольку вектор \mathbf{b}^\perp ортогонален всем векторам из S .
Из (*), если $\mathbf{x} \in \mathbb{R}^n$ — решение системы $A\mathbf{x} = \mathbf{b}_S$, то

$$A^T A\mathbf{x} = A^T \mathbf{b}_S \stackrel{(*)}{=} A^T \mathbf{b}.$$

Обратно, пусть $\mathbf{x} \in \mathbb{R}^n$ является решением системы $A^T A\mathbf{x} = A^T \mathbf{b}$. Возьмем произвольный вектор $\mathbf{y} \in \mathbb{R}^n$ и перемножим вектора $A\mathbf{y}$ и $A\mathbf{x} - \mathbf{b}_S$. Имеем

$$(A\mathbf{y})^T (A\mathbf{x} - \mathbf{b}_S) = \mathbf{y}^T A^T (A\mathbf{x} - \mathbf{b}_S) = \mathbf{y}^T (A^T A\mathbf{x} - A^T \mathbf{b}_S) \stackrel{(*)}{=} \mathbf{y}^T (A^T A\mathbf{x} - A^T \mathbf{b}) = 0.$$

Итак, вектор $A\mathbf{x} - \mathbf{b}_S$ ортогонален любому вектору из S , в частности, самому себе. Отсюда $A\mathbf{x} - \mathbf{b}_S = \mathbf{0}$ и $A\mathbf{x} = \mathbf{b}_S$.

Отметим, что если ранг $k \times n$ -матрицы A равен n , то $n \times n$ -матрица $A^T A$ будет обратимой (*упражнение*). В этом случае при любой правой части $\mathbf{b} \in \mathbb{R}^k$ система $A\mathbf{x} = \mathbf{b}$ имеет единственное псевдорешение, для которого есть простая формула:

$$\mathbf{x} = (A^T A)^{-1} A^T \mathbf{b}.$$

Отметим, что если ранг $k \times n$ -матрицы A равен n , то $n \times n$ -матрица $A^T A$ будет обратимой (*упражнение*). В этом случае при любой правой части $\mathbf{b} \in \mathbb{R}^k$ система $A\mathbf{x} = \mathbf{b}$ имеет единственное псевдорешение, для которого есть простая формула:

$$\mathbf{x} = (A^T A)^{-1} A^T \mathbf{b}.$$

Метод наименьших квадратов изобрел (по его словам — в 1795 г.) и с большим успехом применял Карл Фридрих Гаусс (1777–1855).

Метод наименьших квадратов (3)

Отметим, что если ранг $k \times n$ -матрицы A равен n , то $n \times n$ -матрица $A^T A$ будет обратимой (*упражнение*). В этом случае при любой правой части $\mathbf{b} \in \mathbb{R}^k$ система $A\mathbf{x} = \mathbf{b}$ имеет единственное псевдорешение, для которого есть простая формула:

$$\mathbf{x} = (A^T A)^{-1} A^T \mathbf{b}.$$

Метод наименьших квадратов изобрел (по его словам — в 1795 г.) и с большим успехом применял Карл Фридрих Гаусс (1777–1855).



Метод наименьших квадратов (3)

Отметим, что если ранг $k \times n$ -матрицы A равен n , то $n \times n$ -матрица $A^T A$ будет обратимой (*упражнение*). В этом случае при любой правой части $\mathbf{b} \in \mathbb{R}^k$ система $A\mathbf{x} = \mathbf{b}$ имеет единственное псевдорешение, для которого есть простая формула:

$$\mathbf{x} = (A^T A)^{-1} A^T \mathbf{b}.$$

Метод наименьших квадратов изобрел (по его словам — в 1795 г.) и с большим успехом применял Карл Фридрих Гаусс (1777–1855).



1. Метод наименьших квадратов работает и для унитарных пространств, т.е. для систем с комплексными коэффициентами.

1. Метод наименьших квадратов работает и для унитарных пространств, т.е. для систем с комплексными коэффициентами. Единственное отличие состоит в том, что для отыскания псевдорешений системы линейных уравнений $Ax = b$ надо решать систему $A^*Ax = A^*b$, где A^* — *эрмитово сопряженная* матрица к матрице A . (Эрмитово сопряженная матрица получается, если исходную матрицу транспонировать и заменить каждый элемент его сопряженным: если $A = (a_{ij})_{k \times n}$, то $A^* := (\overline{a_{ji}})_{n \times k}$.)

1. Метод наименьших квадратов работает и для унитарных пространств, т.е. для систем с комплексными коэффициентами. Единственное отличие состоит в том, что для отыскания псевдорешений системы линейных уравнений $Ax = b$ надо решать систему $A^*Ax = A^*b$, где A^* — *эрмитово сопряженная* матрица к матрице A . (Эрмитово сопряженная матрица получается, если исходную матрицу транспонировать и заменить каждый элемент его сопряженным: если $A = (a_{ij})_{k \times n}$, то $A^* := (\overline{a_{ji}})_{n \times k}$.)
2. Имеются и другие методы нахождения псевдорешений несовместных системы линейных уравнений, например, итерационный *метод Качмажа*.

Утверждалось, что если ранг $k \times n$ -матрицы A равен n , то $n \times n$ -матрица $A^T A$ будет обратимой. На самом деле, справедлив более общий факт:

Предложение

Для любой матрицы A над \mathbb{R} ее ранг равен рангу матрицы $A^T A$.

Утверждалось, что если ранг $k \times n$ -матрицы A равен n , то $n \times n$ -матрица $A^T A$ будет обратимой. На самом деле, справедлив более общий факт:

Предложение

Для любой матрицы A над \mathbb{R} ее ранг равен рангу матрицы $A^T A$.

Доказательство. Выше установлено, что системы линейных уравнений $Ax = \mathbf{b}_S$ и $A^T Ax = A^T \mathbf{b}$ равносильны для произвольного вектора \mathbf{b} .

Утверждалось, что если ранг $k \times n$ -матрицы A равен n , то $n \times n$ -матрица $A^T A$ будет обратимой. На самом деле, справедлив более общий факт:

Предложение

Для любой матрицы A над \mathbb{R} ее ранг равен рангу матрицы $A^T A$.

Доказательство. Выше установлено, что системы линейных уравнений $Ax = \mathbf{b}_S$ и $A^T Ax = A^T \mathbf{b}$ равносильны для произвольного вектора \mathbf{b} . Полагая $\mathbf{b} = \mathbf{0}$, заключаем, что у однородных систем $Ax = \mathbf{0}$ и $A^T Ax = \mathbf{0}$ одно и то же пространство решений; обозначим его через R .

Утверждалось, что если ранг $k \times n$ -матрицы A равен n , то $n \times n$ -матрица $A^T A$ будет обратимой. На самом деле, справедлив более общий факт:

Предложение

Для любой матрицы A над \mathbb{R} ее ранг равен рангу матрицы $A^T A$.

Доказательство. Выше установлено, что системы линейных уравнений $Ax = \mathbf{b}_S$ и $A^T Ax = A^T \mathbf{b}$ равносильны для произвольного вектора \mathbf{b} . Полагая $\mathbf{b} = \mathbf{0}$, заключаем, что у однородных систем $Ax = \mathbf{0}$ и $A^T Ax = \mathbf{0}$ одно и то же пространство решений; обозначим его через R . Применяя к каждой из этих систем теорему о размерности пространства решений линейной однородной системы, получаем, что размерность R равна разности между числом неизвестных и рангом матрицы A и в то же время равна разности между числом неизвестных и рангом матрицы $A^T A$.

Утверждалось, что если ранг $k \times n$ -матрицы A равен n , то $n \times n$ -матрица $A^T A$ будет обратимой. На самом деле, справедлив более общий факт:

Предложение

Для любой матрицы A над \mathbb{R} ее ранг равен рангу матрицы $A^T A$.

Доказательство. Выше установлено, что системы линейных уравнений $Ax = \mathbf{b}_S$ и $A^T Ax = A^T \mathbf{b}$ равносильны для произвольного вектора \mathbf{b} . Полагая $\mathbf{b} = \mathbf{0}$, заключаем, что у однородных систем $Ax = \mathbf{0}$ и $A^T Ax = \mathbf{0}$ одно и то же пространство решений; обозначим его через R . Применяя к каждой из этих систем теорему о размерности пространства решений линейной однородной системы, получаем, что размерность R равна разности между числом неизвестных и рангом матрицы A и в то же время равна разности между числом неизвестных и рангом матрицы $A^T A$. Следовательно, эти ранги равны. □

Вопрос

Верен ли аналогичный факт для матриц над произвольными полями?