

Тема II. Линейные операторы

§ 8. Сингулярное разложение

М.В.Волков

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

2021/2022 учебный год

В предыдущем разделе мы обсуждали полезное разложение линейных операторов на пространстве со скалярным произведением (*полярное разложение*). На матричном языке оно отвечает определенному разложению квадратных матриц.

В предыдущем разделе мы обсуждали полезное разложение линейных операторов на пространстве со скалярным произведением (*полярное разложение*). На матричном языке оно отвечает определенному разложению квадратных матриц.

Сейчас рассмотрим ситуацию, когда имеются два пространства со скалярным произведением, U и V , и линейный оператор $\mathcal{A}: U \rightarrow V$.

В предыдущем разделе мы обсуждали полезное разложение линейных операторов на пространстве со скалярным произведением (*полярное разложение*). На матричном языке оно отвечает определенному разложению квадратных матриц.

Сейчас рассмотрим ситуацию, когда имеются два пространства со скалярным произведением, U и V , и линейный оператор $\mathcal{A}: U \rightarrow V$.

Мы укажем некоторое разложение для матриц таких линейных операторов (*сингулярное разложение*). Заметим, что размеры матриц в этом случае произвольны.

В предыдущем разделе мы обсуждали полезное разложение линейных операторов на пространстве со скалярным произведением (*полярное разложение*). На матричном языке оно отвечает определенному разложению квадратных матриц.

Сейчас рассмотрим ситуацию, когда имеются два пространства со скалярным произведением, U и V , и линейный оператор $\mathcal{A}: U \rightarrow V$.

Мы укажем некоторое разложение для матриц таких линейных операторов (*сингулярное разложение*). Заметим, что размеры матриц в этом случае произвольны.

Начнем с одного простого, но очень полезного наблюдения.

Теорема (Эрик Ивар Фредгольм, 1903)

Если $\mathcal{A}: U \rightarrow V$ – линейный оператор пространств со скалярным произведением над полем $F \in \{\mathbb{R}, \mathbb{C}\}$, то $(\text{Im } \mathcal{A})^\perp = \text{Ker } \mathcal{A}^$.*

Теорема (Эрик Ивар Фредгольм, 1903)

Если $\mathcal{A}: U \rightarrow V$ – линейный оператор пространств со скалярным произведением над полем $F \in \{\mathbb{R}, \mathbb{C}\}$, то $(\text{Im } \mathcal{A})^\perp = \text{Ker } \mathcal{A}^*$.

Доказательство. Пусть $y \in \text{Ker } \mathcal{A}^*$, т.е. $\mathcal{A}^*y = 0$.

Теорема (Эрик Ивар Фредгольм, 1903)

Если $\mathcal{A}: U \rightarrow V$ – линейный оператор пространств со скалярным произведением над полем $F \in \{\mathbb{R}, \mathbb{C}\}$, то $(\text{Im } \mathcal{A})^\perp = \text{Ker } \mathcal{A}^*$.

Доказательство. Пусть $y \in \text{Ker } \mathcal{A}^*$, т.е. $\mathcal{A}^*y = 0$. Чтобы доказать, что $y \in (\text{Im } \mathcal{A})^\perp$, возьмем любой вектор $x \in \text{Im } \mathcal{A}$ и проверим, что $y \perp x$.

Теорема (Эрик Ивар Фредгольм, 1903)

Если $\mathcal{A}: U \rightarrow V$ – линейный оператор пространств со скалярным произведением над полем $F \in \{\mathbb{R}, \mathbb{C}\}$, то $(\text{Im } \mathcal{A})^\perp = \text{Ker } \mathcal{A}^*$.

Доказательство. Пусть $y \in \text{Ker } \mathcal{A}^*$, т.е. $\mathcal{A}^*y = 0$. Чтобы доказать, что $y \in (\text{Im } \mathcal{A})^\perp$, возьмем любой вектор $x \in \text{Im } \mathcal{A}$ и проверим, что $y \perp x$. Поскольку $x = \mathcal{A}u$ для некоторого вектора $u \in U$, имеем

$$xy = \mathcal{A}uy = u\mathcal{A}^*y = 0.$$

Теорема (Эрик Ивар Фредгольм, 1903)

Если $\mathcal{A}: U \rightarrow V$ – линейный оператор пространств со скалярным произведением над полем $F \in \{\mathbb{R}, \mathbb{C}\}$, то $(\text{Im } \mathcal{A})^\perp = \text{Ker } \mathcal{A}^*$.

Доказательство. Пусть $\mathbf{y} \in \text{Ker } \mathcal{A}^*$, т.е. $\mathcal{A}^* \mathbf{y} = \mathbf{0}$. Чтобы доказать, что $\mathbf{y} \in (\text{Im } \mathcal{A})^\perp$, возьмем любой вектор $\mathbf{x} \in \text{Im } \mathcal{A}$ и проверим, что $\mathbf{y} \perp \mathbf{x}$. Поскольку $\mathbf{x} = \mathcal{A}\mathbf{u}$ для некоторого вектора $\mathbf{u} \in U$, имеем

$$\mathbf{x}\mathbf{y} = \mathcal{A}\mathbf{u}\mathbf{y} = \mathbf{u}\mathcal{A}^*\mathbf{y} = 0.$$

Обратно, пусть $\mathbf{y} \in (\text{Im } \mathcal{A})^\perp$; тогда $\mathbf{x}\mathbf{y} = 0$ для любого вектора $\mathbf{x} \in \text{Im } \mathcal{A}$.

Теорема (Эрик Ивар Фредгольм, 1903)

Если $\mathcal{A}: U \rightarrow V$ – линейный оператор пространств со скалярным произведением над полем $F \in \{\mathbb{R}, \mathbb{C}\}$, то $(\text{Im } \mathcal{A})^\perp = \text{Ker } \mathcal{A}^*$.

Доказательство. Пусть $y \in \text{Ker } \mathcal{A}^*$, т.е. $\mathcal{A}^*y = 0$. Чтобы доказать, что $y \in (\text{Im } \mathcal{A})^\perp$, возьмем любой вектор $x \in \text{Im } \mathcal{A}$ и проверим, что $y \perp x$. Поскольку $x = \mathcal{A}u$ для некоторого вектора $u \in U$, имеем

$$xy = \mathcal{A}uy = u\mathcal{A}^*y = 0.$$

Обратно, пусть $y \in (\text{Im } \mathcal{A})^\perp$; тогда $xy = 0$ для любого вектора $x \in \text{Im } \mathcal{A}$. Поэтому для произвольного вектора $u \in U$ имеем

$$0 = \mathcal{A}uy = u\mathcal{A}^*y.$$

Теорема (Эрик Ивар Фредгольм, 1903)

Если $\mathcal{A}: U \rightarrow V$ – линейный оператор пространств со скалярным произведением над полем $F \in \{\mathbb{R}, \mathbb{C}\}$, то $(\text{Im } \mathcal{A})^\perp = \text{Ker } \mathcal{A}^*$.

Доказательство. Пусть $y \in \text{Ker } \mathcal{A}^*$, т.е. $\mathcal{A}^*y = 0$. Чтобы доказать, что $y \in (\text{Im } \mathcal{A})^\perp$, возьмем любой вектор $x \in \text{Im } \mathcal{A}$ и проверим, что $y \perp x$. Поскольку $x = \mathcal{A}u$ для некоторого вектора $u \in U$, имеем

$$xy = \mathcal{A}uy = u\mathcal{A}^*y = 0.$$

Обратно, пусть $y \in (\text{Im } \mathcal{A})^\perp$; тогда $xy = 0$ для любого вектора $x \in \text{Im } \mathcal{A}$. Поэтому для произвольного вектора $u \in U$ имеем

$$0 = \mathcal{A}uy = u\mathcal{A}^*y.$$

Вектор \mathcal{A}^*y ортогонален произвольному вектору $u \in U$, и потому он нулевой. Отсюда $y \in \text{Ker } \mathcal{A}^*$. □

Теорема Фредгольма дает ортогональное разложение пространства V :

$$V = \text{Ker } \mathcal{A}^* \oplus \text{Im } \mathcal{A}.$$

Теорема Фредгольма дает ортогональное разложение пространства V :

$$V = \text{Ker } \mathcal{A}^* \oplus \text{Im } \mathcal{A}.$$

Применяя ту же теорему к сопряженному оператору $\mathcal{A}^* : V \rightarrow U$, получим ортогональное разложение пространства U :

$$U = \text{Ker } \mathcal{A} \oplus \text{Im } \mathcal{A}^*.$$

Теорема Фредгольма дает ортогональное разложение пространства V :

$$V = \text{Ker } \mathcal{A}^* \oplus \text{Im } \mathcal{A}.$$

Применяя ту же теорему к сопряженному оператору $\mathcal{A}^* : V \rightarrow U$, получим ортогональное разложение пространства U :

$$U = \text{Ker } \mathcal{A} \oplus \text{Im } \mathcal{A}^*.$$

Положим $U_0 := \text{Im } \mathcal{A}^*$, $V_0 := \text{Im } \mathcal{A}$ и обозначим через \mathcal{A}_0 *ограничение* оператора \mathcal{A} на подпространство U_0 .

Теорема Фредгольма дает ортогональное разложение пространства V :

$$V = \text{Ker } \mathcal{A}^* \oplus \text{Im } \mathcal{A}.$$

Применяя ту же теорему к сопряженному оператору $\mathcal{A}^* : V \rightarrow U$, получим ортогональное разложение пространства U :

$$U = \text{Ker } \mathcal{A} \oplus \text{Im } \mathcal{A}^*.$$

Положим $U_0 := \text{Im } \mathcal{A}^*$, $V_0 := \text{Im } \mathcal{A}$ и обозначим через \mathcal{A}_0 *ограничение* оператора \mathcal{A} на подпространство U_0 . Это означает, что вектор $\mathcal{A}_0 x$ определен, только если $x \in U_0$, и в этом случае $\mathcal{A}_0 x := \mathcal{A}x$.

Теорема Фредгольма дает ортогональное разложение пространства V :

$$V = \text{Ker } \mathcal{A}^* \oplus \text{Im } \mathcal{A}.$$

Применяя ту же теорему к сопряженному оператору $\mathcal{A}^* : V \rightarrow U$, получим ортогональное разложение пространства U :

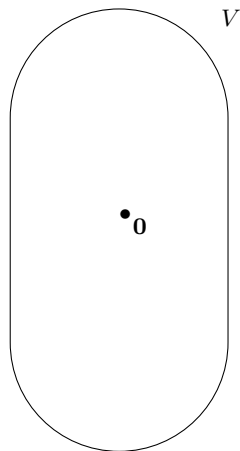
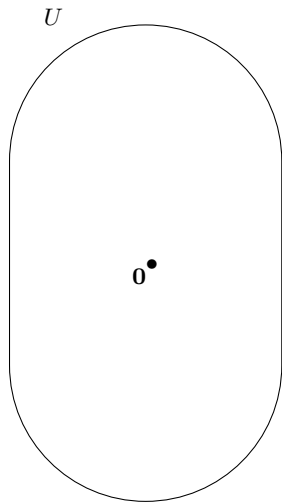
$$U = \text{Ker } \mathcal{A} \oplus \text{Im } \mathcal{A}^*.$$

Положим $U_0 := \text{Im } \mathcal{A}^*$, $V_0 := \text{Im } \mathcal{A}$ и обозначим через \mathcal{A}_0 *ограничение* оператора \mathcal{A} на подпространство U_0 . Это означает, что вектор $\mathcal{A}_0 x$ определен, только если $x \in U_0$, и в этом случае $\mathcal{A}_0 x := \mathcal{A}x$.

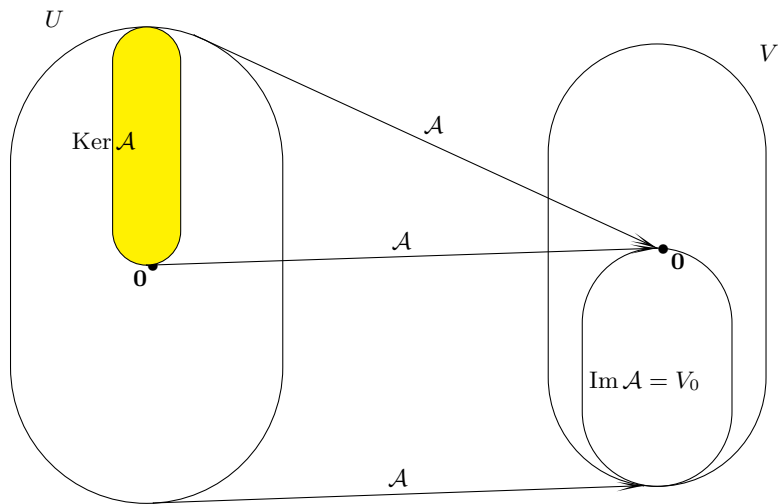
Предложение

\mathcal{A}_0 — изоморфизм пространства U_0 на пространство V_0 .

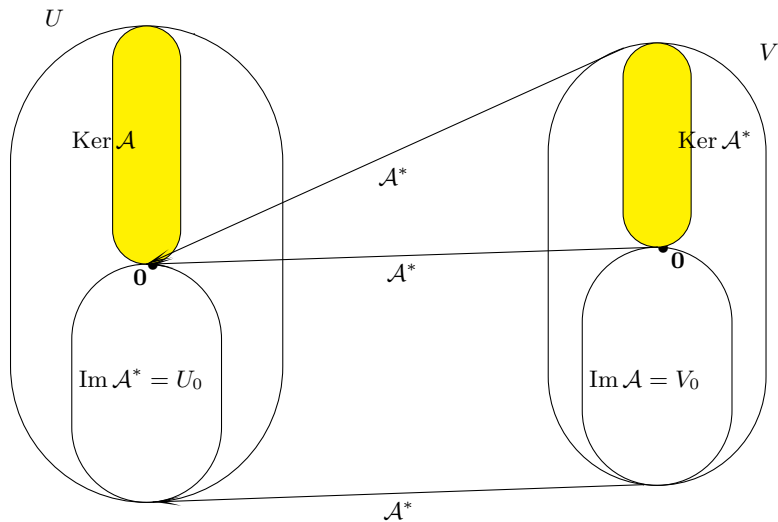
Конфигурация из предложения



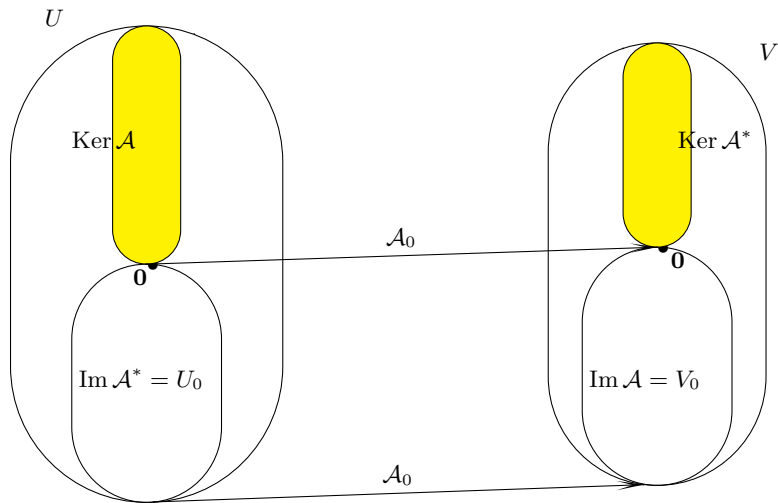
Конфигурация из предложения



Конфигурация из предложения



Конфигурация из предложения



Доказательство. Проверим, что оператор \mathcal{A}_0 взаимно однозначен. Пусть $\mathcal{A}_0 \mathbf{x}_1 = \mathcal{A}_0 \mathbf{x}_2$ для $\mathbf{x}_1, \mathbf{x}_2 \in U_0$.

Доказательство. Проверим, что оператор \mathcal{A}_0 взаимно однозначен. Пусть $\mathcal{A}_0 \mathbf{x}_1 = \mathcal{A}_0 \mathbf{x}_2$ для $\mathbf{x}_1, \mathbf{x}_2 \in U_0$. Тогда $\mathcal{A}_0(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0}$, откуда $\mathbf{x}_1 - \mathbf{x}_2 \in \text{Ker } \mathcal{A}$.

Доказательство. Проверим, что оператор \mathcal{A}_0 взаимно однозначен. Пусть $\mathcal{A}_0 \mathbf{x}_1 = \mathcal{A}_0 \mathbf{x}_2$ для $\mathbf{x}_1, \mathbf{x}_2 \in U_0$. Тогда $\mathcal{A}_0(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0}$, откуда $\mathbf{x}_1 - \mathbf{x}_2 \in \text{Ker } \mathcal{A}$. Но $\text{Ker } \mathcal{A} \cap U_0 = \{\mathbf{0}\}$, поэтому $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$, т.е. $\mathbf{x}_1 = \mathbf{x}_2$.

Доказательство. Проверим, что оператор \mathcal{A}_0 взаимно однозначен. Пусть $\mathcal{A}_0 \mathbf{x}_1 = \mathcal{A}_0 \mathbf{x}_2$ для $\mathbf{x}_1, \mathbf{x}_2 \in U_0$. Тогда $\mathcal{A}_0(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0}$, откуда $\mathbf{x}_1 - \mathbf{x}_2 \in \text{Ker } \mathcal{A}$. Но $\text{Ker } \mathcal{A} \cap U_0 = \{\mathbf{0}\}$, поэтому $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$, т.е. $\mathbf{x}_1 = \mathbf{x}_2$.
Проверим, что \mathcal{A}_0 отображает U_0 на V_0 . Возьмем произвольный $\mathbf{y} \in V_0$. Поскольку $V_0 = \text{Im } \mathcal{A}$, найдется вектор $\mathbf{x} \in U$ такой, что $\mathbf{y} = \mathcal{A}\mathbf{x}$.

Доказательство. Проверим, что оператор \mathcal{A}_0 взаимно однозначен. Пусть $\mathcal{A}_0 \mathbf{x}_1 = \mathcal{A}_0 \mathbf{x}_2$ для $\mathbf{x}_1, \mathbf{x}_2 \in U_0$. Тогда $\mathcal{A}_0(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0}$, откуда $\mathbf{x}_1 - \mathbf{x}_2 \in \text{Ker } \mathcal{A}$. Но $\text{Ker } \mathcal{A} \cap U_0 = \{\mathbf{0}\}$, поэтому $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$, т.е. $\mathbf{x}_1 = \mathbf{x}_2$.

Проверим, что \mathcal{A}_0 отображает U_0 на V_0 . Возьмем произвольный $\mathbf{y} \in V_0$. Поскольку $V_0 = \text{Im } \mathcal{A}$, найдется вектор $\mathbf{x} \in U$ такой, что $\mathbf{y} = \mathcal{A}\mathbf{x}$. Пользуясь ортогональным разложением $U = \text{Ker } \mathcal{A} \oplus U_0$, представим \mathbf{x} как $\mathbf{x} = \mathbf{x}_0 + \mathbf{z}$, где $\mathbf{x}_0 \in U_0$, а $\mathbf{z} \in \text{Ker } \mathcal{A}$.

Доказательство. Проверим, что оператор \mathcal{A}_0 взаимно однозначен. Пусть $\mathcal{A}_0 \mathbf{x}_1 = \mathcal{A}_0 \mathbf{x}_2$ для $\mathbf{x}_1, \mathbf{x}_2 \in U_0$. Тогда $\mathcal{A}_0(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0}$, откуда $\mathbf{x}_1 - \mathbf{x}_2 \in \text{Ker } \mathcal{A}$. Но $\text{Ker } \mathcal{A} \cap U_0 = \{\mathbf{0}\}$, поэтому $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$, т.е. $\mathbf{x}_1 = \mathbf{x}_2$.

Проверим, что \mathcal{A}_0 отображает U_0 на V_0 . Возьмем произвольный $\mathbf{y} \in V_0$. Поскольку $V_0 = \text{Im } \mathcal{A}$, найдется вектор $\mathbf{x} \in U$ такой, что $\mathbf{y} = \mathcal{A}\mathbf{x}$. Пользуясь ортогональным разложением $U = \text{Ker } \mathcal{A} \oplus U_0$, представим \mathbf{x} как $\mathbf{x} = \mathbf{x}_0 + \mathbf{z}$, где $\mathbf{x}_0 \in U_0$, а $\mathbf{z} \in \text{Ker } \mathcal{A}$. Тогда

$$\mathbf{y} = \mathcal{A}\mathbf{x} = \mathcal{A}(\mathbf{x}_0 + \mathbf{z}) = \mathcal{A}\mathbf{x}_0 + \mathcal{A}\mathbf{z}$$

Доказательство. Проверим, что оператор \mathcal{A}_0 взаимно однозначен. Пусть $\mathcal{A}_0 \mathbf{x}_1 = \mathcal{A}_0 \mathbf{x}_2$ для $\mathbf{x}_1, \mathbf{x}_2 \in U_0$. Тогда $\mathcal{A}_0(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0}$, откуда $\mathbf{x}_1 - \mathbf{x}_2 \in \text{Ker } \mathcal{A}$. Но $\text{Ker } \mathcal{A} \cap U_0 = \{\mathbf{0}\}$, поэтому $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$, т.е. $\mathbf{x}_1 = \mathbf{x}_2$.

Проверим, что \mathcal{A}_0 отображает U_0 на V_0 . Возьмем произвольный $\mathbf{y} \in V_0$. Поскольку $V_0 = \text{Im } \mathcal{A}$, найдется вектор $\mathbf{x} \in U$ такой, что $\mathbf{y} = \mathcal{A}\mathbf{x}$. Пользуясь ортогональным разложением $U = \text{Ker } \mathcal{A} \oplus U_0$, представим \mathbf{x} как $\mathbf{x} = \mathbf{x}_0 + \mathbf{z}$, где $\mathbf{x}_0 \in U_0$, а $\mathbf{z} \in \text{Ker } \mathcal{A}$. Тогда

$$\mathbf{y} = \mathcal{A}\mathbf{x} = \mathcal{A}(\mathbf{x}_0 + \mathbf{z}) = \mathcal{A}\mathbf{x}_0 + \mathcal{A}\mathbf{z} = \mathcal{A}\mathbf{x}_0 = \mathcal{A}_0\mathbf{x}_0. \quad \square$$

Доказательство. Проверим, что оператор \mathcal{A}_0 взаимно однозначен. Пусть $\mathcal{A}_0 \mathbf{x}_1 = \mathcal{A}_0 \mathbf{x}_2$ для $\mathbf{x}_1, \mathbf{x}_2 \in U_0$. Тогда $\mathcal{A}_0(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0}$, откуда $\mathbf{x}_1 - \mathbf{x}_2 \in \text{Ker } \mathcal{A}$. Но $\text{Ker } \mathcal{A} \cap U_0 = \{\mathbf{0}\}$, поэтому $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$, т.е. $\mathbf{x}_1 = \mathbf{x}_2$.

Проверим, что \mathcal{A}_0 отображает U_0 на V_0 . Возьмем произвольный $\mathbf{y} \in V_0$. Поскольку $V_0 = \text{Im } \mathcal{A}$, найдется вектор $\mathbf{x} \in U$ такой, что $\mathbf{y} = \mathcal{A}\mathbf{x}$. Пользуясь ортогональным разложением $U = \text{Ker } \mathcal{A} \oplus U_0$, представим \mathbf{x} как $\mathbf{x} = \mathbf{x}_0 + \mathbf{z}$, где $\mathbf{x}_0 \in U_0$, а $\mathbf{z} \in \text{Ker } \mathcal{A}$. Тогда

$$\mathbf{y} = \mathcal{A}\mathbf{x} = \mathcal{A}(\mathbf{x}_0 + \mathbf{z}) = \mathcal{A}\mathbf{x}_0 + \mathcal{A}\mathbf{z} = \mathcal{A}\mathbf{x}_0 = \mathcal{A}_0\mathbf{x}_0. \quad \square$$

В силу предложения $\dim U_0 = \dim V_0$; эта размерность равна рангу оператора \mathcal{A} , который мы обозначим через r .

Доказательство. Проверим, что оператор \mathcal{A}_0 взаимно однозначен. Пусть $\mathcal{A}_0 \mathbf{x}_1 = \mathcal{A}_0 \mathbf{x}_2$ для $\mathbf{x}_1, \mathbf{x}_2 \in U_0$. Тогда $\mathcal{A}_0(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0}$, откуда $\mathbf{x}_1 - \mathbf{x}_2 \in \text{Ker } \mathcal{A}$. Но $\text{Ker } \mathcal{A} \cap U_0 = \{\mathbf{0}\}$, поэтому $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$, т.е. $\mathbf{x}_1 = \mathbf{x}_2$.

Проверим, что \mathcal{A}_0 отображает U_0 на V_0 . Возьмем произвольный $\mathbf{y} \in V_0$. Поскольку $V_0 = \text{Im } \mathcal{A}$, найдется вектор $\mathbf{x} \in U$ такой, что $\mathbf{y} = \mathcal{A}\mathbf{x}$. Пользуясь ортогональным разложением $U = \text{Ker } \mathcal{A} \oplus U_0$, представим \mathbf{x} как $\mathbf{x} = \mathbf{x}_0 + \mathbf{z}$, где $\mathbf{x}_0 \in U_0$, а $\mathbf{z} \in \text{Ker } \mathcal{A}$. Тогда

$$\mathbf{y} = \mathcal{A}\mathbf{x} = \mathcal{A}(\mathbf{x}_0 + \mathbf{z}) = \mathcal{A}\mathbf{x}_0 + \mathcal{A}\mathbf{z} = \mathcal{A}\mathbf{x}_0 = \mathcal{A}_0\mathbf{x}_0. \quad \square$$

В силу предложения $\dim U_0 = \dim V_0$; эта размерность равна рангу оператора \mathcal{A} , который мы обозначим через r . Построим «согласованные» ортонормированные базисы в U_0 и V_0 .

∇1. $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{A}\mathcal{A}^*$.

∇1. $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{A}\mathcal{A}^*$.

Доказательство. Ясно, что $\text{Ker } \mathcal{A} \subseteq \text{Ker } \mathcal{A}\mathcal{A}^*$.

$\nabla 1.$ $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{A}\mathcal{A}^*.$

Доказательство. Ясно, что $\text{Ker } \mathcal{A} \subseteq \text{Ker } \mathcal{A}\mathcal{A}^*.$ Обратно, пусть $x \in \text{Ker } \mathcal{A}\mathcal{A}^*,$ т.е. $(\mathcal{A}\mathcal{A}^*)x = \mathbf{0}.$

∇1. $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{A}\mathcal{A}^*$.

Доказательство. Ясно, что $\text{Ker } \mathcal{A} \subseteq \text{Ker } \mathcal{A}\mathcal{A}^*$. Обратно, пусть $\mathbf{x} \in \text{Ker } \mathcal{A}\mathcal{A}^*$, т.е. $(\mathcal{A}\mathcal{A}^*)\mathbf{x} = \mathbf{0}$. Тогда

$$0 = \mathbf{x}((\mathcal{A}\mathcal{A}^*)\mathbf{x}) = \mathbf{x}(\mathcal{A}^*(\mathcal{A}\mathbf{x}))$$

∇1. $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{A}\mathcal{A}^*$.

Доказательство. Ясно, что $\text{Ker } \mathcal{A} \subseteq \text{Ker } \mathcal{A}\mathcal{A}^*$. Обратно, пусть $\mathbf{x} \in \text{Ker } \mathcal{A}\mathcal{A}^*$, т.е. $(\mathcal{A}\mathcal{A}^*)\mathbf{x} = \mathbf{0}$. Тогда

$$0 = \mathbf{x}((\mathcal{A}\mathcal{A}^*)\mathbf{x}) = \mathbf{x}(\mathcal{A}^*(\mathcal{A}\mathbf{x})) = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}).$$

∇1. $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{A}\mathcal{A}^*$.

Доказательство. Ясно, что $\text{Ker } \mathcal{A} \subseteq \text{Ker } \mathcal{A}\mathcal{A}^*$. Обратно, пусть $\mathbf{x} \in \text{Ker } \mathcal{A}\mathcal{A}^*$, т.е. $(\mathcal{A}\mathcal{A}^*)\mathbf{x} = \mathbf{0}$. Тогда

$$0 = \mathbf{x}((\mathcal{A}\mathcal{A}^*)\mathbf{x}) = \mathbf{x}(\mathcal{A}^*(\mathcal{A}\mathbf{x})) = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}).$$

Отсюда $\mathcal{A}\mathbf{x} = \mathbf{0}$, т.е. $\mathbf{x} \in \text{Ker } \mathcal{A}$. □

∇1. $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{A}\mathcal{A}^*$.

Доказательство. Ясно, что $\text{Ker } \mathcal{A} \subseteq \text{Ker } \mathcal{A}\mathcal{A}^*$. Обратно, пусть $\mathbf{x} \in \text{Ker } \mathcal{A}\mathcal{A}^*$, т.е. $(\mathcal{A}\mathcal{A}^*)\mathbf{x} = \mathbf{0}$. Тогда

$$0 = \mathbf{x}((\mathcal{A}\mathcal{A}^*)\mathbf{x}) = \mathbf{x}(\mathcal{A}^*(\mathcal{A}\mathbf{x})) = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}).$$

Отсюда $\mathcal{A}\mathbf{x} = \mathbf{0}$, т.е. $\mathbf{x} \in \text{Ker } \mathcal{A}$. □

Из ∇1 следует, что ограничение оператора $\mathcal{A}\mathcal{A}^*$ на подпространство U_0 невырождено.

∇1. $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{A}\mathcal{A}^*$.

Доказательство. Ясно, что $\text{Ker } \mathcal{A} \subseteq \text{Ker } \mathcal{A}\mathcal{A}^*$. Обратно, пусть $x \in \text{Ker } \mathcal{A}\mathcal{A}^*$, т.е. $(\mathcal{A}\mathcal{A}^*)x = \mathbf{0}$. Тогда

$$0 = x((\mathcal{A}\mathcal{A}^*)x) = x(\mathcal{A}^*(\mathcal{A}x)) = (\mathcal{A}x)(\mathcal{A}x).$$

Отсюда $\mathcal{A}x = \mathbf{0}$, т.е. $x \in \text{Ker } \mathcal{A}$. □

Из ∇1 следует, что ограничение оператора $\mathcal{A}\mathcal{A}^*$ на подпространство U_0 невырождено. Тогда это ограничение – положительный оператор, так как $((\mathcal{A}\mathcal{A}^*)x)x$

∇1. $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{A}\mathcal{A}^*$.

Доказательство. Ясно, что $\text{Ker } \mathcal{A} \subseteq \text{Ker } \mathcal{A}\mathcal{A}^*$. Обратно, пусть $\mathbf{x} \in \text{Ker } \mathcal{A}\mathcal{A}^*$, т.е. $(\mathcal{A}\mathcal{A}^*)\mathbf{x} = \mathbf{0}$. Тогда

$$0 = \mathbf{x}((\mathcal{A}\mathcal{A}^*)\mathbf{x}) = \mathbf{x}(\mathcal{A}^*(\mathcal{A}\mathbf{x})) = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}).$$

Отсюда $\mathcal{A}\mathbf{x} = \mathbf{0}$, т.е. $\mathbf{x} \in \text{Ker } \mathcal{A}$. □

Из ∇1 следует, что ограничение оператора $\mathcal{A}\mathcal{A}^*$ на подпространство U_0 невырождено. Тогда это ограничение – положительный оператор, так как $((\mathcal{A}\mathcal{A}^*)\mathbf{x})\mathbf{x} = (\mathcal{A}^*(\mathcal{A}\mathbf{x}))\mathbf{x}$

∇1. $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{A}\mathcal{A}^*$.

Доказательство. Ясно, что $\text{Ker } \mathcal{A} \subseteq \text{Ker } \mathcal{A}\mathcal{A}^*$. Обратно, пусть $\mathbf{x} \in \text{Ker } \mathcal{A}\mathcal{A}^*$, т.е. $(\mathcal{A}\mathcal{A}^*)\mathbf{x} = \mathbf{0}$. Тогда

$$0 = \mathbf{x}((\mathcal{A}\mathcal{A}^*)\mathbf{x}) = \mathbf{x}(\mathcal{A}^*(\mathcal{A}\mathbf{x})) = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}).$$

Отсюда $\mathcal{A}\mathbf{x} = \mathbf{0}$, т.е. $\mathbf{x} \in \text{Ker } \mathcal{A}$. □

Из ∇1 следует, что ограничение оператора $\mathcal{A}\mathcal{A}^*$ на подпространство U_0 невырождено. Тогда это ограничение – положительный оператор, так как $((\mathcal{A}\mathcal{A}^*)\mathbf{x})\mathbf{x} = (\mathcal{A}^*(\mathcal{A}\mathbf{x}))\mathbf{x} = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}) > 0$ для ненулевых векторов $\mathbf{x} \in U_0$.

∇1. $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{A}\mathcal{A}^*$.

Доказательство. Ясно, что $\text{Ker } \mathcal{A} \subseteq \text{Ker } \mathcal{A}\mathcal{A}^*$. Обратно, пусть $\mathbf{x} \in \text{Ker } \mathcal{A}\mathcal{A}^*$, т.е. $(\mathcal{A}\mathcal{A}^*)\mathbf{x} = \mathbf{0}$. Тогда

$$0 = \mathbf{x}((\mathcal{A}\mathcal{A}^*)\mathbf{x}) = \mathbf{x}(\mathcal{A}^*(\mathcal{A}\mathbf{x})) = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}).$$

Отсюда $\mathcal{A}\mathbf{x} = \mathbf{0}$, т.е. $\mathbf{x} \in \text{Ker } \mathcal{A}$. □

Из ∇1 следует, что ограничение оператора $\mathcal{A}\mathcal{A}^*$ на подпространство U_0 невырождено. Тогда это ограничение – положительный оператор, так как $((\mathcal{A}\mathcal{A}^*)\mathbf{x})\mathbf{x} = (\mathcal{A}^*(\mathcal{A}\mathbf{x}))\mathbf{x} = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}) > 0$ для ненулевых векторов $\mathbf{x} \in U_0$.

Выберем в U_0 ортонормированный базис $\mathbf{u}_1, \dots, \mathbf{u}_r$ из собственных векторов оператора $\mathcal{A}\mathcal{A}^*$, принадлежащих собственным значениям $\lambda_1, \dots, \lambda_r > 0$.

∇1. $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{A}\mathcal{A}^*$.

Доказательство. Ясно, что $\text{Ker } \mathcal{A} \subseteq \text{Ker } \mathcal{A}\mathcal{A}^*$. Обратно, пусть $\mathbf{x} \in \text{Ker } \mathcal{A}\mathcal{A}^*$, т.е. $(\mathcal{A}\mathcal{A}^*)\mathbf{x} = \mathbf{0}$. Тогда

$$0 = \mathbf{x}((\mathcal{A}\mathcal{A}^*)\mathbf{x}) = \mathbf{x}(\mathcal{A}^*(\mathcal{A}\mathbf{x})) = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}).$$

Отсюда $\mathcal{A}\mathbf{x} = \mathbf{0}$, т.е. $\mathbf{x} \in \text{Ker } \mathcal{A}$. □

Из ∇1 следует, что ограничение оператора $\mathcal{A}\mathcal{A}^*$ на подпространство U_0 невырождено. Тогда это ограничение – положительный оператор, так как $((\mathcal{A}\mathcal{A}^*)\mathbf{x})\mathbf{x} = (\mathcal{A}^*(\mathcal{A}\mathbf{x}))\mathbf{x} = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}) > 0$ для ненулевых векторов $\mathbf{x} \in U_0$.

Выберем в U_0 ортонормированный базис $\mathbf{u}_1, \dots, \mathbf{u}_r$ из собственных векторов оператора $\mathcal{A}\mathcal{A}^*$, принадлежащих собственным значениям $\lambda_1, \dots, \lambda_r > 0$. Образы $\mathcal{A}\mathbf{u}_1, \dots, \mathcal{A}\mathbf{u}_r$ этих векторов лежат в V_0 и попарно ортогональны.

∇1. $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{A}\mathcal{A}^*$.

Доказательство. Ясно, что $\text{Ker } \mathcal{A} \subseteq \text{Ker } \mathcal{A}\mathcal{A}^*$. Обратно, пусть $\mathbf{x} \in \text{Ker } \mathcal{A}\mathcal{A}^*$, т.е. $(\mathcal{A}\mathcal{A}^*)\mathbf{x} = \mathbf{0}$. Тогда

$$0 = \mathbf{x}((\mathcal{A}\mathcal{A}^*)\mathbf{x}) = \mathbf{x}(\mathcal{A}^*(\mathcal{A}\mathbf{x})) = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}).$$

Отсюда $\mathcal{A}\mathbf{x} = \mathbf{0}$, т.е. $\mathbf{x} \in \text{Ker } \mathcal{A}$. □

Из ∇1 следует, что ограничение оператора $\mathcal{A}\mathcal{A}^*$ на подпространство U_0 невырождено. Тогда это ограничение – положительный оператор, так как $((\mathcal{A}\mathcal{A}^*)\mathbf{x})\mathbf{x} = (\mathcal{A}^*(\mathcal{A}\mathbf{x}))\mathbf{x} = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}) > 0$ для ненулевых векторов $\mathbf{x} \in U_0$.

Выберем в U_0 ортонормированный базис $\mathbf{u}_1, \dots, \mathbf{u}_r$ из собственных векторов оператора $\mathcal{A}\mathcal{A}^*$, принадлежащих собственным значениям $\lambda_1, \dots, \lambda_r > 0$. Образы $\mathcal{A}\mathbf{u}_1, \dots, \mathcal{A}\mathbf{u}_r$ этих векторов лежат в V_0 и попарно ортогональны. Действительно, при $i \neq j$

$$\mathcal{A}\mathbf{u}_i\mathcal{A}\mathbf{u}_j = \mathbf{u}_i(\mathcal{A}^*(\mathcal{A}\mathbf{u}_j)) = \mathbf{u}_i((\mathcal{A}\mathcal{A}^*)\mathbf{u}_j)$$

∇1. $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{A}\mathcal{A}^*$.

Доказательство. Ясно, что $\text{Ker } \mathcal{A} \subseteq \text{Ker } \mathcal{A}\mathcal{A}^*$. Обратно, пусть $\mathbf{x} \in \text{Ker } \mathcal{A}\mathcal{A}^*$, т.е. $(\mathcal{A}\mathcal{A}^*)\mathbf{x} = \mathbf{0}$. Тогда

$$0 = \mathbf{x}((\mathcal{A}\mathcal{A}^*)\mathbf{x}) = \mathbf{x}(\mathcal{A}^*(\mathcal{A}\mathbf{x})) = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}).$$

Отсюда $\mathcal{A}\mathbf{x} = \mathbf{0}$, т.е. $\mathbf{x} \in \text{Ker } \mathcal{A}$. □

Из ∇1 следует, что ограничение оператора $\mathcal{A}\mathcal{A}^*$ на подпространство U_0 невырождено. Тогда это ограничение – положительный оператор, так как $((\mathcal{A}\mathcal{A}^*)\mathbf{x})\mathbf{x} = (\mathcal{A}^*(\mathcal{A}\mathbf{x}))\mathbf{x} = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}) > 0$ для ненулевых векторов $\mathbf{x} \in U_0$.

Выберем в U_0 ортонормированный базис $\mathbf{u}_1, \dots, \mathbf{u}_r$ из собственных векторов оператора $\mathcal{A}\mathcal{A}^*$, принадлежащих собственным значениям $\lambda_1, \dots, \lambda_r > 0$. Образы $\mathcal{A}\mathbf{u}_1, \dots, \mathcal{A}\mathbf{u}_r$ этих векторов лежат в V_0 и попарно ортогональны. Действительно, при $i \neq j$

$$\mathcal{A}\mathbf{u}_i\mathcal{A}\mathbf{u}_j = \mathbf{u}_i(\mathcal{A}^*(\mathcal{A}\mathbf{u}_j)) = \mathbf{u}_i((\mathcal{A}\mathcal{A}^*)\mathbf{u}_j) = \mathbf{u}_i(\lambda_j\mathbf{u}_j) = \lambda_j\mathbf{u}_i\mathbf{u}_j = 0.$$

∇1. $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{A}\mathcal{A}^*$.

Доказательство. Ясно, что $\text{Ker } \mathcal{A} \subseteq \text{Ker } \mathcal{A}\mathcal{A}^*$. Обратно, пусть $\mathbf{x} \in \text{Ker } \mathcal{A}\mathcal{A}^*$, т.е. $(\mathcal{A}\mathcal{A}^*)\mathbf{x} = \mathbf{0}$. Тогда

$$0 = \mathbf{x}((\mathcal{A}\mathcal{A}^*)\mathbf{x}) = \mathbf{x}(\mathcal{A}^*(\mathcal{A}\mathbf{x})) = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}).$$

Отсюда $\mathcal{A}\mathbf{x} = \mathbf{0}$, т.е. $\mathbf{x} \in \text{Ker } \mathcal{A}$. □

Из ∇1 следует, что ограничение оператора $\mathcal{A}\mathcal{A}^*$ на подпространство U_0 невырождено. Тогда это ограничение – положительный оператор, так как $((\mathcal{A}\mathcal{A}^*)\mathbf{x})\mathbf{x} = (\mathcal{A}^*(\mathcal{A}\mathbf{x}))\mathbf{x} = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}) > 0$ для ненулевых векторов $\mathbf{x} \in U_0$.

Выберем в U_0 ортонормированный базис $\mathbf{u}_1, \dots, \mathbf{u}_r$ из собственных векторов оператора $\mathcal{A}\mathcal{A}^*$, принадлежащих собственным значениям $\lambda_1, \dots, \lambda_r > 0$. Образы $\mathcal{A}\mathbf{u}_1, \dots, \mathcal{A}\mathbf{u}_r$ этих векторов лежат в V_0 и попарно ортогональны. Действительно, при $i \neq j$

$$\mathcal{A}\mathbf{u}_i\mathcal{A}\mathbf{u}_j = \mathbf{u}_i(\mathcal{A}^*(\mathcal{A}\mathbf{u}_j)) = \mathbf{u}_i((\mathcal{A}\mathcal{A}^*)\mathbf{u}_j) = \mathbf{u}_i(\lambda_j\mathbf{u}_j) = \lambda_j\mathbf{u}_i\mathbf{u}_j = 0.$$

Положим $\mu_i := \sqrt{\lambda_i}$ и $\mathbf{v}_i := \mu_i^{-1}\mathcal{A}\mathbf{u}_i$ для каждого $i = 1, \dots, r$.

∇1. $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{A}\mathcal{A}^*$.

Доказательство. Ясно, что $\text{Ker } \mathcal{A} \subseteq \text{Ker } \mathcal{A}\mathcal{A}^*$. Обратно, пусть $\mathbf{x} \in \text{Ker } \mathcal{A}\mathcal{A}^*$, т.е. $(\mathcal{A}\mathcal{A}^*)\mathbf{x} = \mathbf{0}$. Тогда

$$0 = \mathbf{x}((\mathcal{A}\mathcal{A}^*)\mathbf{x}) = \mathbf{x}(\mathcal{A}^*(\mathcal{A}\mathbf{x})) = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}).$$

Отсюда $\mathcal{A}\mathbf{x} = \mathbf{0}$, т.е. $\mathbf{x} \in \text{Ker } \mathcal{A}$. □

Из ∇1 следует, что ограничение оператора $\mathcal{A}\mathcal{A}^*$ на подпространство U_0 невырождено. Тогда это ограничение – положительный оператор, так как $((\mathcal{A}\mathcal{A}^*)\mathbf{x})\mathbf{x} = (\mathcal{A}^*(\mathcal{A}\mathbf{x}))\mathbf{x} = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}) > 0$ для ненулевых векторов $\mathbf{x} \in U_0$.

Выберем в U_0 ортонормированный базис $\mathbf{u}_1, \dots, \mathbf{u}_r$ из собственных векторов оператора $\mathcal{A}\mathcal{A}^*$, принадлежащих собственным значениям $\lambda_1, \dots, \lambda_r > 0$. Образы $\mathcal{A}\mathbf{u}_1, \dots, \mathcal{A}\mathbf{u}_r$ этих векторов лежат в V_0 и попарно ортогональны. Действительно, при $i \neq j$

$$\mathcal{A}\mathbf{u}_i\mathcal{A}\mathbf{u}_j = \mathbf{u}_i(\mathcal{A}^*(\mathcal{A}\mathbf{u}_j)) = \mathbf{u}_i((\mathcal{A}\mathcal{A}^*)\mathbf{u}_j) = \mathbf{u}_i(\lambda_j\mathbf{u}_j) = \lambda_j\mathbf{u}_i\mathbf{u}_j = 0.$$

Положим $\mu_i := \sqrt{\lambda_i}$ и $\mathbf{v}_i := \mu_i^{-1}\mathcal{A}\mathbf{u}_i$ для каждого $i = 1, \dots, r$. Тогда

$$|\mathbf{v}_i|^2 = \mathbf{v}_i\mathbf{v}_i = \frac{(\mathcal{A}\mathbf{u}_i)(\mathcal{A}\mathbf{u}_i)}{\mu_i^2}$$

∇1. $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{A}\mathcal{A}^*$.

Доказательство. Ясно, что $\text{Ker } \mathcal{A} \subseteq \text{Ker } \mathcal{A}\mathcal{A}^*$. Обратно, пусть $\mathbf{x} \in \text{Ker } \mathcal{A}\mathcal{A}^*$, т.е. $(\mathcal{A}\mathcal{A}^*)\mathbf{x} = \mathbf{0}$. Тогда

$$0 = \mathbf{x}((\mathcal{A}\mathcal{A}^*)\mathbf{x}) = \mathbf{x}(\mathcal{A}^*(\mathcal{A}\mathbf{x})) = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}).$$

Отсюда $\mathcal{A}\mathbf{x} = \mathbf{0}$, т.е. $\mathbf{x} \in \text{Ker } \mathcal{A}$. □

Из ∇1 следует, что ограничение оператора $\mathcal{A}\mathcal{A}^*$ на подпространство U_0 невырождено. Тогда это ограничение – положительный оператор, так как $((\mathcal{A}\mathcal{A}^*)\mathbf{x})\mathbf{x} = (\mathcal{A}^*(\mathcal{A}\mathbf{x}))\mathbf{x} = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}) > 0$ для ненулевых векторов $\mathbf{x} \in U_0$.

Выберем в U_0 ортонормированный базис $\mathbf{u}_1, \dots, \mathbf{u}_r$ из собственных векторов оператора $\mathcal{A}\mathcal{A}^*$, принадлежащих собственным значениям $\lambda_1, \dots, \lambda_r > 0$. Образы $\mathcal{A}\mathbf{u}_1, \dots, \mathcal{A}\mathbf{u}_r$ этих векторов лежат в V_0 и попарно ортогональны. Действительно, при $i \neq j$

$$\mathcal{A}\mathbf{u}_i\mathcal{A}\mathbf{u}_j = \mathbf{u}_i(\mathcal{A}^*(\mathcal{A}\mathbf{u}_j)) = \mathbf{u}_i((\mathcal{A}\mathcal{A}^*)\mathbf{u}_j) = \mathbf{u}_i(\lambda_j\mathbf{u}_j) = \lambda_j\mathbf{u}_i\mathbf{u}_j = 0.$$

Положим $\mu_i := \sqrt{\lambda_i}$ и $\mathbf{v}_i := \mu_i^{-1}\mathcal{A}\mathbf{u}_i$ для каждого $i = 1, \dots, r$. Тогда

$$|\mathbf{v}_i|^2 = \mathbf{v}_i\mathbf{v}_i = \frac{(\mathcal{A}\mathbf{u}_i)(\mathcal{A}\mathbf{u}_i)}{\mu_i^2} = \frac{\mathbf{u}_i(\mathcal{A}^*(\mathcal{A}\mathbf{u}_i))}{\mu_i^2} = \frac{\mathbf{u}_i((\mathcal{A}\mathcal{A}^*)\mathbf{u}_i)}{\mu_i^2}$$

∇1. $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{A}\mathcal{A}^*$.

Доказательство. Ясно, что $\text{Ker } \mathcal{A} \subseteq \text{Ker } \mathcal{A}\mathcal{A}^*$. Обратно, пусть $\mathbf{x} \in \text{Ker } \mathcal{A}\mathcal{A}^*$, т.е. $(\mathcal{A}\mathcal{A}^*)\mathbf{x} = \mathbf{0}$. Тогда

$$0 = \mathbf{x}((\mathcal{A}\mathcal{A}^*)\mathbf{x}) = \mathbf{x}(\mathcal{A}^*(\mathcal{A}\mathbf{x})) = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}).$$

Отсюда $\mathcal{A}\mathbf{x} = \mathbf{0}$, т.е. $\mathbf{x} \in \text{Ker } \mathcal{A}$. □

Из ∇1 следует, что ограничение оператора $\mathcal{A}\mathcal{A}^*$ на подпространство U_0 невырождено. Тогда это ограничение – положительный оператор, так как $((\mathcal{A}\mathcal{A}^*)\mathbf{x})\mathbf{x} = (\mathcal{A}^*(\mathcal{A}\mathbf{x}))\mathbf{x} = (\mathcal{A}\mathbf{x})(\mathcal{A}\mathbf{x}) > 0$ для ненулевых векторов $\mathbf{x} \in U_0$.

Выберем в U_0 ортонормированный базис $\mathbf{u}_1, \dots, \mathbf{u}_r$ из собственных векторов оператора $\mathcal{A}\mathcal{A}^*$, принадлежащих собственным значениям $\lambda_1, \dots, \lambda_r > 0$. Образы $\mathcal{A}\mathbf{u}_1, \dots, \mathcal{A}\mathbf{u}_r$ этих векторов лежат в V_0 и попарно ортогональны. Действительно, при $i \neq j$

$$\mathcal{A}\mathbf{u}_i\mathcal{A}\mathbf{u}_j = \mathbf{u}_i(\mathcal{A}^*(\mathcal{A}\mathbf{u}_j)) = \mathbf{u}_i((\mathcal{A}\mathcal{A}^*)\mathbf{u}_j) = \mathbf{u}_i(\lambda_j\mathbf{u}_j) = \lambda_j\mathbf{u}_i\mathbf{u}_j = 0.$$

Положим $\mu_i := \sqrt{\lambda_i}$ и $\mathbf{v}_i := \mu_i^{-1}\mathcal{A}\mathbf{u}_i$ для каждого $i = 1, \dots, r$. Тогда

$$|\mathbf{v}_i|^2 = \mathbf{v}_i\mathbf{v}_i = \frac{(\mathcal{A}\mathbf{u}_i)(\mathcal{A}\mathbf{u}_i)}{\mu_i^2} = \frac{\mathbf{u}_i(\mathcal{A}^*(\mathcal{A}\mathbf{u}_i))}{\mu_i^2} = \frac{\mathbf{u}_i((\mathcal{A}\mathcal{A}^*)\mathbf{u}_i)}{\mu_i^2} = \frac{\lambda_i}{\mu_i^2} = 1.$$

Итак, вектора $\mathbf{v}_1, \dots, \mathbf{v}_r$ образуют ортонормированный базис подпространства V_0 .

Итак, вектора $\mathbf{v}_1, \dots, \mathbf{v}_r$ образуют ортонормированный базис подпространства V_0 . Дополним систему $\mathbf{v}_1, \dots, \mathbf{v}_r$ ортонормированным базисом пространства $\text{Ker } \mathcal{A}^* = (\text{Im } \mathcal{A})^\perp$ до ортонормированного базиса $\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ всего пространства V , здесь $n := \dim V$.

Итак, вектора $\mathbf{v}_1, \dots, \mathbf{v}_r$ образуют ортонормированный базис подпространства V_0 . Дополним систему $\mathbf{v}_1, \dots, \mathbf{v}_r$ ортонормированным базисом пространства $\text{Ker } \mathcal{A}^* = (\text{Im } \mathcal{A})^\perp$ до ортонормированного базиса $\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ всего пространства V , здесь $n := \dim V$. Аналогично, систему $\mathbf{u}_1, \dots, \mathbf{u}_r$ дополним ортонормированным базисом пространства $\text{Ker } \mathcal{A} = (\text{Im } \mathcal{A}^*)^\perp$ до ортонормированного базиса $\mathbf{u}_1, \dots, \mathbf{u}_r, \mathbf{u}_{r+1}, \dots, \mathbf{u}_k$ всего пространства U , здесь $k := \dim U$.

Итак, вектора $\mathbf{v}_1, \dots, \mathbf{v}_r$ образуют ортонормированный базис подпространства V_0 . Дополним систему $\mathbf{v}_1, \dots, \mathbf{v}_r$ ортонормированным базисом пространства $\text{Ker } \mathcal{A}^* = (\text{Im } \mathcal{A})^\perp$ до ортонормированного базиса $\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ всего пространства V , здесь $n := \dim V$. Аналогично, систему $\mathbf{u}_1, \dots, \mathbf{u}_r$ дополним ортонормированным базисом пространства $\text{Ker } \mathcal{A} = (\text{Im } \mathcal{A}^*)^\perp$ до ортонормированного базиса $\mathbf{u}_1, \dots, \mathbf{u}_r, \mathbf{u}_{r+1}, \dots, \mathbf{u}_k$ всего пространства U , здесь $k := \dim U$. Эти ортонормированные базисы называются *сингулярными базисами* оператора \mathcal{A} , а числа $\mu_1, \mu_2, \dots, \mu_r$ – его *сингулярными числами*.

Итак, вектора $\mathbf{v}_1, \dots, \mathbf{v}_r$ образуют ортонормированный базис подпространства V_0 . Дополним систему $\mathbf{v}_1, \dots, \mathbf{v}_r$ ортонормированным базисом пространства $\text{Ker } \mathcal{A}^* = (\text{Im } \mathcal{A})^\perp$ до ортонормированного базиса $\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ всего пространства V , здесь $n := \dim V$. Аналогично, систему $\mathbf{u}_1, \dots, \mathbf{u}_r$ дополним ортонормированным базисом пространства $\text{Ker } \mathcal{A} = (\text{Im } \mathcal{A}^*)^\perp$ до ортонормированного базиса $\mathbf{u}_1, \dots, \mathbf{u}_r, \mathbf{u}_{r+1}, \dots, \mathbf{u}_k$ всего пространства U , здесь $k := \dim U$. Эти ортонормированные базисы называются *сингулярными базисами* оператора \mathcal{A} , а числа $\mu_1, \mu_2, \dots, \mu_r$ – его *сингулярными числами*. Обычно первые r векторов в сингулярных базисах нумеруют так, чтобы сингулярные числа шли в порядке убывания: $\mu_1 \geq \mu_2 \geq \dots \geq \mu_r > 0$.

Итак, вектора $\mathbf{v}_1, \dots, \mathbf{v}_r$ образуют ортонормированный базис подпространства V_0 . Дополним систему $\mathbf{v}_1, \dots, \mathbf{v}_r$ ортонормированным базисом пространства $\text{Ker } \mathcal{A}^* = (\text{Im } \mathcal{A})^\perp$ до ортонормированного базиса $\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ всего пространства V , здесь $n := \dim V$. Аналогично, систему $\mathbf{u}_1, \dots, \mathbf{u}_r$ дополним ортонормированным базисом пространства $\text{Ker } \mathcal{A} = (\text{Im } \mathcal{A}^*)^\perp$ до ортонормированного базиса $\mathbf{u}_1, \dots, \mathbf{u}_r, \mathbf{u}_{r+1}, \dots, \mathbf{u}_k$ всего пространства U , здесь $k := \dim U$. Эти ортонормированные базисы называются *сингулярными базисами* оператора \mathcal{A} , а числа $\mu_1, \mu_2, \dots, \mu_r$ – его *сингулярными числами*. Обычно первые r векторов в сингулярных базисах нумеруют так, чтобы сингулярные числа шли в порядке убывания: $\mu_1 \geq \mu_2 \geq \dots \geq \mu_r > 0$. Действие оператора \mathcal{A} в его сингулярных базисах прозрачно донельзя:

$$\mathcal{A}\mathbf{u}_i = \begin{cases} \mu_i \mathbf{v}_i, & \text{если } i \leq r, \\ \mathbf{0}, & \text{если } i > r. \end{cases}$$

Итак, вектора $\mathbf{v}_1, \dots, \mathbf{v}_r$ образуют ортонормированный базис подпространства V_0 . Дополним систему $\mathbf{v}_1, \dots, \mathbf{v}_r$ ортонормированным базисом пространства $\text{Ker } \mathcal{A}^* = (\text{Im } \mathcal{A})^\perp$ до ортонормированного базиса $\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ всего пространства V , здесь $n := \dim V$. Аналогично, систему $\mathbf{u}_1, \dots, \mathbf{u}_r$ дополним ортонормированным базисом пространства $\text{Ker } \mathcal{A} = (\text{Im } \mathcal{A}^*)^\perp$ до ортонормированного базиса $\mathbf{u}_1, \dots, \mathbf{u}_r, \mathbf{u}_{r+1}, \dots, \mathbf{u}_k$ всего пространства U , здесь $k := \dim U$. Эти ортонормированные базисы называются *сингулярными базисами* оператора \mathcal{A} , а числа $\mu_1, \mu_2, \dots, \mu_r$ – его *сингулярными числами*. Обычно первые r векторов в сингулярных базисах нумеруют так, чтобы сингулярные числа шли в порядке убывания: $\mu_1 \geq \mu_2 \geq \dots \geq \mu_r > 0$. Действие оператора \mathcal{A} в его сингулярных базисах прозрачно донельзя:

$$\mathcal{A}\mathbf{u}_i = \begin{cases} \mu_i \mathbf{v}_i, & \text{если } i \leq r, \\ \mathbf{0}, & \text{если } i > r. \end{cases}$$

Поэтому матрица оператора \mathcal{A} в его сингулярных базисах устроена так: в ней на местах $(1, 1), (2, 2), \dots, (r, r)$ стоят соответствующие сингулярные числа $\mu_1, \mu_2, \dots, \mu_r$, а на всех остальных местах – нули.

Мы доказали бóльшую часть следующего результата:

Теорема (сингулярное представление линейного оператора)

Для любого линейного оператора $A: U \rightarrow V$ пространств со скалярным произведением в U и V можно выбрать ортонормированные базисы, в которых его матрица имеет вид

$$A = \begin{pmatrix} \mu_1 & 0 & \dots & 0 & 0 & \dots \\ 0 & \mu_2 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \mu_r & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \end{pmatrix}, \quad (\#)$$

где r – ранг A , а $\mu_1, \mu_2, \dots, \mu_r$ – положительные действительные числа, определяемые однозначно с точностью до порядка.

Мы доказали бóльшую часть следующего результата:

Теорема (сингулярное представление линейного оператора)

Для любого линейного оператора $A: U \rightarrow V$ пространств со скалярным произведением в U и V можно выбрать ортонормированные базисы, в которых его матрица имеет вид

$$A = \begin{pmatrix} \mu_1 & 0 & \dots & 0 & 0 & \dots \\ 0 & \mu_2 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \mu_r & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \end{pmatrix}, \quad (\#)$$

где r – ранг A , а $\mu_1, \mu_2, \dots, \mu_r$ – положительные действительные числа, определяемые однозначно с точностью до порядка.

Доказательство. Осталось доказать только то, что если матрица оператора A в каких-то ортонормированных базисах имеет вид $(\#)$, то диагональные элементы этой матрицы определены однозначно с точностью до порядка.

Действительно, если B_U и B_V – какие-то ортонормированные базисы в U и V , а A – матрица оператора \mathcal{A} в этих базисах, то матрица сопряженного оператора \mathcal{A}^* в тех же базисах есть A^* .

Действительно, если B_U и B_V – какие-то ортонормированные базисы в U и V , а A – матрица оператора \mathcal{A} в этих базисах, то матрица сопряженного оператора \mathcal{A}^* в тех же базисах есть A^* . Если A имеет вид (‡), то $A^* = A^T$.

Действительно, если B_U и B_V – какие-то ортонормированные базисы в U и V , а A – матрица оператора \mathcal{A} в этих базисах, то матрица сопряженного оператора \mathcal{A}^* в тех же базисах есть A^* . Если A имеет вид (‡), то $A^* = A^T$. Поэтому матрица оператора $\mathcal{A}\mathcal{A}^*$ в базисе B_U – квадратная диагональная матрица с ненулевыми элементами диагонали, равными $\mu_1^2, \mu_2^2, \dots, \mu_r^2$.

Действительно, если B_U и B_V – какие-то ортонормированные базисы в U и V , а A – матрица оператора \mathcal{A} в этих базисах, то матрица сопряженного оператора \mathcal{A}^* в тех же базисах есть A^* . Если A имеет вид (‡), то $A^* = A^T$. Поэтому матрица оператора $\mathcal{A}\mathcal{A}^*$ в базисе B_U – квадратная диагональная матрица с ненулевыми элементами диагонали, равными $\mu_1^2, \mu_2^2, \dots, \mu_r^2$. Если матрица оператора в некотором базисе диагональна, то на диагонали стоят его собственные значения.

Действительно, если B_U и B_V – какие-то ортонормированные базисы в U и V , а A – матрица оператора \mathcal{A} в этих базисах, то матрица сопряженного оператора \mathcal{A}^* в тех же базисах есть A^* . Если A имеет вид (‡), то $A^* = A^T$. Поэтому матрица оператора $\mathcal{A}\mathcal{A}^*$ в базисе B_U – квадратная диагональная матрица с ненулевыми элементами диагонали, равными $\mu_1^2, \mu_2^2, \dots, \mu_r^2$. Если матрица оператора в некотором базисе диагональна, то на диагонали стоят его собственные значения. Значит, $\mu_1^2, \mu_2^2, \dots, \mu_r^2$ – собственные значения оператора $\mathcal{A}\mathcal{A}^*$, а следовательно, $\mu_1, \mu_2, \dots, \mu_r$ суть в точности сингулярные числа оператора \mathcal{A} и не зависят от базисов B_U и B_V . \square

Следствие (SVD, Джеймс Джозеф Сильвестр, 1889)

Пусть F – одно из полей \mathbb{R} или \mathbb{C} , а M – произвольная $k \times n$ -матрица над F . Существует пара ортогональных (унитарных) матриц $R \in M_k(F)$ и $S \in M_n(F)$ такая, что $M = RAS$, где A – матрица вида $\begin{pmatrix} \mu_1 & & & \\ & \mu_2 & & \\ & & \dots & \\ & & & \mu_r & & \\ & & & & & 0 \end{pmatrix}$, где r – ранг матрицы M , а $\mu_1, \mu_2, \dots, \mu_r$ – положительные действительные числа, определяемые однозначно с точностью до порядка.

Следствие (SVD, Джеймс Джозеф Сильвестр, 1889)

Пусть F – одно из полей \mathbb{R} или \mathbb{C} , а M – произвольная $k \times n$ -матрица над F . Существует пара ортогональных (унитарных) матриц $R \in M_k(F)$ и $S \in M_n(F)$ такая, что $M = RAS$, где A – матрица вида $\begin{pmatrix} \mu_1 & & & \\ & \mu_2 & & \\ & & \dots & \\ & & & \mu_r & & \\ & & & & & 0 \end{pmatrix}$, где r – ранг матрицы M , а $\mu_1, \mu_2, \dots, \mu_r$ – положительные действительные числа, определяемые однозначно с точностью до порядка.

Доказательство. Определим оператор $A: F^n \rightarrow F^k$ таким правилом:
 $Ax := Mx$.

Следствие (SVD, Джеймс Джозеф Сильвестр, 1889)

Пусть F – одно из полей \mathbb{R} или \mathbb{C} , а M – произвольная $k \times n$ -матрица над F . Существует пара ортогональных (унитарных) матриц $R \in M_k(F)$ и $S \in M_n(F)$ такая, что $M = RAS$, где A – матрица вида $(\#)$, где r – ранг матрицы M , а $\mu_1, \mu_2, \dots, \mu_r$ – положительные действительные числа, определяемые однозначно с точностью до порядка.

Доказательство. Определим оператор $\mathcal{A}: F^n \rightarrow F^k$ таким правилом: $\mathcal{A}x := Mx$. По теореме в F^n и F^k есть ортонормированные базисы, скажем, B_n и B_k , в которых матрица A оператора \mathcal{A} имеет вид $(\#)$.

Следствие (SVD, Джеймс Джозеф Сильвестр, 1889)

Пусть F – одно из полей \mathbb{R} или \mathbb{C} , а M – произвольная $k \times n$ -матрица над F . Существует пара ортогональных (унитарных) матриц $R \in M_k(F)$ и $S \in M_n(F)$ такая, что $M = RAS$, где A – матрица вида $(\#)$, где r – ранг матрицы M , а $\mu_1, \mu_2, \dots, \mu_r$ – положительные действительные числа, определяемые однозначно с точностью до порядка.

Доказательство. Определим оператор $\mathcal{A}: F^n \rightarrow F^k$ таким правилом: $\mathcal{A}x := Mx$. По теореме в F^n и F^k есть ортонормированные базисы, скажем, B_n и B_k , в которых матрица A оператора \mathcal{A} имеет вид $(\#)$. С другой стороны, M есть матрица оператора \mathcal{A} в стандартных ортонормированных базисах пространств F^n и F^k .

Следствие (SVD, Джеймс Джозеф Сильвестр, 1889)

Пусть F – одно из полей \mathbb{R} или \mathbb{C} , а M – произвольная $k \times n$ -матрица над F . Существует пара ортогональных (унитарных) матриц $R \in M_k(F)$ и $S \in M_n(F)$ такая, что $M = RAS$, где A – матрица вида $(\#)$, где r – ранг матрицы M , а $\mu_1, \mu_2, \dots, \mu_r$ – положительные действительные числа, определяемые однозначно с точностью до порядка.

Доказательство. Определим оператор $\mathcal{A}: F^n \rightarrow F^k$ таким правилом: $\mathcal{A}x := Mx$. По теореме в F^n и F^k есть ортонормированные базисы, скажем, B_n и B_k , в которых матрица A оператора \mathcal{A} имеет вид $(\#)$. С другой стороны, M есть матрица оператора \mathcal{A} в стандартных ортонормированных базисах пространств F^n и F^k . Обозначим через R матрицу перехода от стандартного базиса пространства F^k к базису B_k , а через S матрицу перехода от базиса B_n к стандартному базису пространства F^n .

Следствие (SVD, Джеймс Джозеф Сильвестр, 1889)

Пусть F – одно из полей \mathbb{R} или \mathbb{C} , а M – произвольная $k \times n$ -матрица над F . Существует пара ортогональных (унитарных) матриц $R \in M_k(F)$ и $S \in M_n(F)$ такая, что $M = RAS$, где A – матрица вида $(\#)$, где r – ранг матрицы M , а $\mu_1, \mu_2, \dots, \mu_r$ – положительные действительные числа, определяемые однозначно с точностью до порядка.

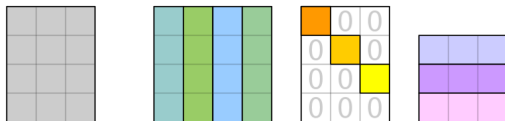
Доказательство. Определим оператор $\mathcal{A}: F^n \rightarrow F^k$ таким правилом: $\mathcal{A}x := Mx$. По теореме в F^n и F^k есть ортонормированные базисы, скажем, B_n и B_k , в которых матрица A оператора \mathcal{A} имеет вид $(\#)$. С другой стороны, M есть матрица оператора \mathcal{A} в стандартных ортонормированных базисах пространств F^n и F^k . Обозначим через R матрицу перехода от стандартного базиса пространства F^k к базису B_k , а через S матрицу перехода от базиса B_n к стандартному базису пространства F^n . Тогда матрицы R и S ортогональные (унитарные), и по формуле замены матрицы оператора $M = RAS$. □

Следствие (SVD, Джеймс Джозеф Сильвестр, 1889)

Пусть F – одно из полей \mathbb{R} или \mathbb{C} , а M – произвольная $k \times n$ -матрица над F . Существует пара ортогональных (унитарных) матриц $R \in M_k(F)$ и $S \in M_n(F)$ такая, что $M = RAS$, где A – матрица вида $(\#)$, где r – ранг матрицы M , а $\mu_1, \mu_2, \dots, \mu_r$ – положительные действительные числа, определяемые однозначно с точностью до порядка.

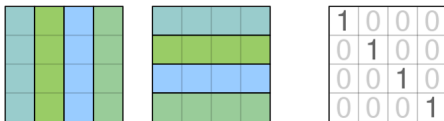
Доказательство. Определим оператор $\mathcal{A}: F^n \rightarrow F^k$ таким правилом: $\mathcal{A}x := Mx$. По теореме в F^n и F^k есть ортонормированные базисы, скажем, B_n и B_k , в которых матрица A оператора \mathcal{A} имеет вид $(\#)$. С другой стороны, M есть матрица оператора \mathcal{A} в стандартных ортонормированных базисах пространств F^n и F^k . Обозначим через R матрицу перехода от стандартного базиса пространства F^k к базису B_k , а через S матрицу перехода от базиса B_n к стандартному базису пространства F^n . Тогда матрицы R и S ортогональные (унитарные), и по формуле замены матрицы оператора $M = RAS$. □

Замечание. Речь идет о формуле $A_{P_2, Q_2} = T_{Q_2 \rightarrow Q_1} A_{P_1, Q_1} T_{P_1 \rightarrow P_2}$ из §II.1.

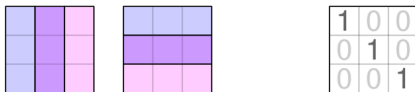


$$\mathbf{M} = \mathbf{R} \mathbf{A} \mathbf{S}$$

$m \times n$ $m \times m$ $m \times n$ $n \times n$



$$\mathbf{R}^* = \mathbf{E}_m$$



$$\mathbf{S}^* = \mathbf{E}_n$$

Полярное разложение из предыдущей лекции легко получается из SVD.

Полярное разложение из предыдущей лекции легко получается из SVD. Действительно, пусть M – произвольная квадратная матрица над одним из полей \mathbb{R} или \mathbb{C} .

Полярное разложение из предыдущей лекции легко получается из SVD. Действительно, пусть M – произвольная квадратная матрица над одним из полей \mathbb{R} или \mathbb{C} . По SVD имеем $M = RAS$, где A – матрица вида $(\#)$, а R и S – ортогональные (унитарные) матрицы.

Полярное разложение из предыдущей лекции легко получается из SVD.

Действительно, пусть M – произвольная квадратная матрица над одним из полей \mathbb{R} или \mathbb{C} . По SVD имеем $M = RAS$, где A – матрица вида $\begin{pmatrix} \# \\ & 0 \end{pmatrix}$, а R и S – ортогональные (унитарные) матрицы. Тогда $M = (RS)(S^{-1}AS)$, матрица RS ортогональна (унитарна), а $S^{-1}AS$ – неотрицательная симметрическая (эрмитова) матрица.

Полярное разложение из предыдущей лекции легко получается из SVD.

Действительно, пусть M – произвольная квадратная матрица над одним из полей \mathbb{R} или \mathbb{C} . По SVD имеем $M = RAS$, где A – матрица вида $\begin{pmatrix} \# \\ & 0 \end{pmatrix}$, а R и S – ортогональные (унитарные) матрицы. Тогда $M = (RS)(S^{-1}AS)$, матрица RS ортогональна (унитарна), а $S^{-1}AS$ – неотрицательная симметрическая (эрмитова) матрица.

Итак, каждая квадратная действительная (комплексная) матрица представима в виде произведения ортогональной (унитарной) и симметрической (эрмитовой) матриц.

Полярное разложение из предыдущей лекции легко получается из SVD.

Действительно, пусть M – произвольная квадратная матрица над одним из полей \mathbb{R} или \mathbb{C} . По SVD имеем $M = RAS$, где A – матрица вида $\begin{pmatrix} \# \\ & 0 \end{pmatrix}$, а R и S – ортогональные (унитарные) матрицы. Тогда $M = (RS)(S^{-1}AS)$, матрица RS ортогональна (унитарна), а $S^{-1}AS$ – неотрицательная симметрическая (эрмитова) матрица.

Итак, каждая квадратная действительная (комплексная) матрица представима в виде произведения ортогональной (унитарной) и симметрической (эрмитовой) матриц.

На операторном языке это означает, что любой линейный оператор \mathcal{A} на евклидовом (унитарном) пространстве представим в виде $\mathcal{A} = \mathcal{B}\mathcal{U}$, где \mathcal{B} – неотрицательный оператор, а \mathcal{U} – ортогональный (унитарный) оператор.

Напомним, что *псевдорешение* системы линейных уравнений $Ax = b$ — это вектор x_0 , минимизирующий расстояние между векторами Ax и b .

Напомним, что *псевдорешение* системы линейных уравнений $Ax = b$ — это вектор x_0 , минимизирующий расстояние между векторами Ax и b .

В осеннем семестре было доказано, что псевдорешения суть в точности решения (в обычном смысле) системы $A^*Ax = A^*b$, где A^* — *эрмитово сопряженная* матрица к A (метод наименьших квадратов). Эта система всегда совместна.

Напомним, что *псевдорешение* системы линейных уравнений $Ax = b$ — это вектор x_0 , минимизирующий расстояние между векторами Ax и b .

В осеннем семестре было доказано, что псевдорешения суть в точности решения (в обычном смысле) системы $A^*Ax = A^*b$, где A^* — *эрмитово сопряженная* матрица к A (метод наименьших квадратов). Эта система всегда совместна.

Если ранг матрицы A равен числу неизвестных, система $A^*Ax = A^*b$ имеет единственное решение, а следовательно, исходная система $Ax = b$ имеет *единственное* псевдорешение.

Напомним, что *псевдорешение* системы линейных уравнений $Ax = b$ — это вектор x_0 , минимизирующий расстояние между векторами Ax и b .

В осеннем семестре было доказано, что псевдорешения суть в точности решения (в обычном смысле) системы $A^*Ax = A^*b$, где A^* — *эрмитово сопряженная* матрица к A (метод наименьших квадратов). Эта система всегда совместна.

Если ранг матрицы A равен числу неизвестных, система $A^*Ax = A^*b$ имеет единственное решение, а следовательно, исходная система $Ax = b$ имеет *единственное* псевдорешение.

Если псевдорешение неединственно, то интересуются псевдорешением наименьшей длины (оно называется *нормальным* псевдорешением).

Напомним, что *псевдорешение* системы линейных уравнений $Ax = b$ — это вектор x_0 , минимизирующий расстояние между векторами Ax и b .

В осеннем семестре было доказано, что псевдорешения суть в точности решения (в обычном смысле) системы $A^*Ax = A^*b$, где A^* — *эрмитово сопряженная* матрица к A (метод наименьших квадратов). Эта система всегда совместна.

Если ранг матрицы A равен числу неизвестных, система $A^*Ax = A^*b$ имеет единственное решение, а следовательно, исходная система $Ax = b$ имеет *единственное* псевдорешение.

Если псевдорешение неединственно, то интересуются псевдорешением наименьшей длины (оно называется *нормальным* псевдорешением).

Займемся вопросом о том, как находить нормальные псевдорешения.

Напомним, что *псевдорешение* системы линейных уравнений $Ax = b$ — это вектор x_0 , минимизирующий расстояние между векторами Ax и b .

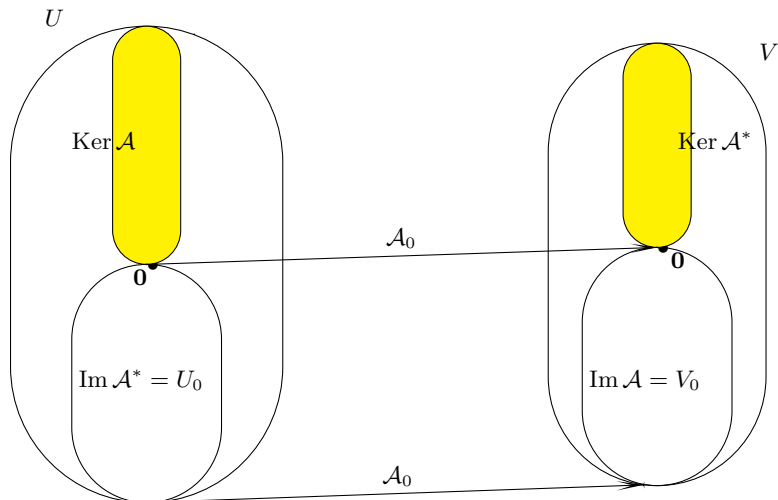
В осеннем семестре было доказано, что псевдорешения суть в точности решения (в обычном смысле) системы $A^*Ax = A^*b$, где A^* — *эрмитово сопряженная* матрица к A (метод наименьших квадратов). Эта система всегда совместна.

Если ранг матрицы A равен числу неизвестных, система $A^*Ax = A^*b$ имеет единственное решение, а следовательно, исходная система $Ax = b$ имеет *единственное* псевдорешение.

Если псевдорешение неединственно, то интересуются псевдорешением наименьшей длины (оно называется *нормальным* псевдорешением).

Займемся вопросом о том, как находить нормальные псевдорешения. Заметим, что этот вопрос представляет интерес и для совместных систем с бесконечным множеством решений (*недоопределенных систем*).

Рассмотрим снова конфигурацию из предложения об операторе \mathcal{A}_0 , который взаимно однозначно отображает $U_0 = \text{Im } \mathcal{A}^*$ на $V_0 = \text{Im } \mathcal{A}$:



У отображения $\mathcal{A}_0: U_0 \rightarrow V_0$ есть обратное отображение $\mathcal{A}_0^{-1}: V_0 \rightarrow U_0$.

У отображения $\mathcal{A}_0: U_0 \rightarrow V_0$ есть обратное отображение $\mathcal{A}_0^{-1}: V_0 \rightarrow U_0$.
Обозначим через \mathcal{P}_0 ортопроектор V на V_0 и положим $\mathcal{A}^+ := \mathcal{A}_0^{-1}\mathcal{P}_0$.

У отображения $\mathcal{A}_0: U_0 \rightarrow V_0$ есть обратное отображение $\mathcal{A}_0^{-1}: V_0 \rightarrow U_0$. Обозначим через \mathcal{P}_0 ортопроектор V на V_0 и положим $\mathcal{A}^+ := \mathcal{A}_0^{-1}\mathcal{P}_0$. Линейное отображение $\mathcal{A}^+: V \rightarrow U$ называется *псевдообратным* к отображению \mathcal{A} (или *обратным Мура–Пенроуза*).

У отображения $\mathcal{A}_0: U_0 \rightarrow V_0$ есть обратное отображение $\mathcal{A}_0^{-1}: V_0 \rightarrow U_0$. Обозначим через \mathcal{P}_0 ортопроектор V на V_0 и положим $\mathcal{A}^+ := \mathcal{A}_0^{-1}\mathcal{P}_0$. Линейное отображение $\mathcal{A}^+: V \rightarrow U$ называется *псевдообратным* к отображению \mathcal{A} (или *обратным Мура–Пенроуза*).

Идея псевдообращения восходит к работе Эрика Ивара Фредгольма 1903 г. Элиаким Гастингс Мур ввел это понятие в 1920 г., Арне Бьерхаммар — в 1951 г., а Роджер Пенроуз (нобелевский лауреат 2020 года) — в 1955 г.

У отображения $\mathcal{A}_0: U_0 \rightarrow V_0$ есть обратное отображение $\mathcal{A}_0^{-1}: V_0 \rightarrow U_0$. Обозначим через \mathcal{P}_0 ортопроектор V на V_0 и положим $\mathcal{A}^+ := \mathcal{A}_0^{-1}\mathcal{P}_0$. Линейное отображение $\mathcal{A}^+: V \rightarrow U$ называется *псевдообратным* к отображению \mathcal{A} (или *обратным Мура–Пенроуза*).

Идея псевдообращения восходит к работе Эрика Ивара Фредгольма 1903 г. Элиаким Гастингс Мур ввел это понятие в 1920 г., Арне Бьерхаммар — в 1951 г., а Роджер Пенроуз (нобелевский лауреат 2020 года) — в 1955 г.

Если A – матрица линейного отображения $\mathcal{A}: U \rightarrow V$ в некоторых ортонормированных базисах пространств U и V , то матрица псевдообратного отображения $\mathcal{A}^+: V \rightarrow U$ в тех же базисах называется *псевдообратной* к матрице A и обозначается через A^+ .

У отображения $\mathcal{A}_0: U_0 \rightarrow V_0$ есть обратное отображение $\mathcal{A}_0^{-1}: V_0 \rightarrow U_0$. Обозначим через \mathcal{P}_0 ортопроектор V на V_0 и положим $\mathcal{A}^+ := \mathcal{A}_0^{-1}\mathcal{P}_0$. Линейное отображение $\mathcal{A}^+: V \rightarrow U$ называется *псевдообратным* к отображению \mathcal{A} (или *обратным Мура–Пенроуза*).

Идея псевдообращения восходит к работе Эрика Ивара Фредгольма 1903 г. Элиаким Гастингс Мур ввел это понятие в 1920 г., Арне Бьерхаммар — в 1951 г., а Роджер Пенроуз (нобелевский лауреат 2020 года) — в 1955 г.

Если A – матрица линейного отображения $\mathcal{A}: U \rightarrow V$ в некоторых ортонормированных базисах пространств U и V , то матрица псевдообратного отображения $\mathcal{A}^+: V \rightarrow U$ в тех же базисах называется *псевдообратной* к матрице A и обозначается через A^+ . Отметим, что псевдообратная матрица существует для любой матрицы (не обязательно квадратной). Если же матрица A квадратная и обратимая, то $A^+ = A^{-1}$.

Теорема (нормальное псевдорешение)

Для любой системы линейных уравнений $Ax = \mathbf{b}$ над полем $F \in \{\mathbb{R}, \mathbb{C}\}$ формула $\mathbf{x} = A^+\mathbf{b}$ возвращает ее нормальное псевдорешение.

Теорема (нормальное псевдорешение)

Для любой системы линейных уравнений $Ax = \mathbf{b}$ над полем $F \in \{\mathbb{R}, \mathbb{C}\}$ формула $\mathbf{x} = A^+ \mathbf{b}$ возвращает ее нормальное псевдорешение.

Доказательство. Пусть A — $n \times k$ -матрица. Введем линейное отображение $\mathcal{A}: F^k \rightarrow F^n$, определяемое матрицей A по правилу $\mathcal{A}\mathbf{u} := A\mathbf{u}$; тогда матрица A^+ отвечает отображению \mathcal{A}^+ .

Теорема (нормальное псевдорешение)

Для любой системы линейных уравнений $Ax = \mathbf{b}$ над полем $F \in \{\mathbb{R}, \mathbb{C}\}$ формула $\mathbf{x} = A^+ \mathbf{b}$ возвращает ее нормальное псевдорешение.

Доказательство. Пусть A — $n \times k$ -матрица. Введем линейное отображение $\mathcal{A}: F^k \rightarrow F^n$, определяемое матрицей A по правилу $\mathcal{A}\mathbf{u} := A\mathbf{u}$; тогда матрица A^+ отвечает отображению \mathcal{A}^+ . По построению $A^+ = A_0^{-1}P_0$, где A_0 и P_0 — матрицы обратимого отображения \mathcal{A}_0 и ортопроектора \mathcal{P}_0 соответственно.

Теорема (нормальное псевдорешение)

Для любой системы линейных уравнений $Ax = \mathbf{b}$ над полем $F \in \{\mathbb{R}, \mathbb{C}\}$ формула $\mathbf{x} = A^+ \mathbf{b}$ возвращает ее нормальное псевдорешение.

Доказательство. Пусть A — $n \times k$ -матрица. Введем линейное отображение $\mathcal{A}: F^k \rightarrow F^n$, определяемое матрицей A по правилу $\mathcal{A}\mathbf{u} := A\mathbf{u}$; тогда матрица A^+ отвечает отображению \mathcal{A}^+ . По построению $A^+ = A_0^{-1}P_0$, где A_0 и P_0 — матрицы обратимого отображения \mathcal{A}_0 и ортопроектора \mathcal{P}_0 соответственно. $P_0 \mathbf{b}$ — это ортогональная проекция столбца \mathbf{b} на подпространство $\text{Im } \mathcal{A}$, поэтому псевдорешения системы $Ax = \mathbf{b}$ — это столбцы $\mathbf{x} \in F^k$, удовлетворяющие $Ax = P_0 \mathbf{b}$.

Теорема (нормальное псевдорешение)

Для любой системы линейных уравнений $Ax = \mathbf{b}$ над полем $F \in \{\mathbb{R}, \mathbb{C}\}$ формула $\mathbf{x} = A^+\mathbf{b}$ возвращает ее нормальное псевдорешение.

Доказательство. Пусть A — $n \times k$ -матрица. Введем линейное отображение $\mathcal{A}: F^k \rightarrow F^n$, определяемое матрицей A по правилу $\mathcal{A}\mathbf{u} := A\mathbf{u}$; тогда матрица A^+ отвечает отображению \mathcal{A}^+ . По построению $A^+ = A_0^{-1}P_0$, где A_0 и P_0 — матрицы обратимого отображения \mathcal{A}_0 и ортопроектора \mathcal{P}_0 соответственно. $P_0\mathbf{b}$ — это ортогональная проекция столбца \mathbf{b} на подпространство $\text{Im } \mathcal{A}$, поэтому псевдорешения системы $Ax = \mathbf{b}$ — это столбцы $\mathbf{x} \in F^k$, удовлетворяющие $Ax = P_0\mathbf{b}$. Столбец $A^+\mathbf{b} \in \text{Im } \mathcal{A}^*$ этому равенству удовлетворяет, так как $A(A^+\mathbf{b}) = A_0(A_0^{-1}P_0\mathbf{b}) = P_0\mathbf{b}$.

Теорема (нормальное псевдорешение)

Для любой системы линейных уравнений $Ax = \mathbf{b}$ над полем $F \in \{\mathbb{R}, \mathbb{C}\}$ формула $\mathbf{x} = A^+\mathbf{b}$ возвращает ее нормальное псевдорешение.

Доказательство. Пусть A — $n \times k$ -матрица. Введем линейное отображение $\mathcal{A}: F^k \rightarrow F^n$, определяемое матрицей A по правилу $\mathcal{A}\mathbf{u} := A\mathbf{u}$; тогда матрица A^+ отвечает отображению \mathcal{A}^+ . По построению $A^+ = A_0^{-1}P_0$, где A_0 и P_0 — матрицы обратимого отображения \mathcal{A}_0 и ортопроектора \mathcal{P}_0 соответственно. $P_0\mathbf{b}$ — это ортогональная проекция столбца \mathbf{b} на подпространство $\text{Im } \mathcal{A}$, поэтому псевдорешения системы $Ax = \mathbf{b}$ — это столбцы $\mathbf{x} \in F^k$, удовлетворяющие $Ax = P_0\mathbf{b}$. Столбец $A^+\mathbf{b} \in \text{Im } \mathcal{A}^*$ этому равенству удовлетворяет, так как $A(A^+\mathbf{b}) = A_0(A_0^{-1}P_0\mathbf{b}) = P_0\mathbf{b}$. Итак, $A^+\mathbf{b}$ — псевдорешение, и любое псевдорешение представимо в виде $\mathbf{x} = A^+\mathbf{b} + \mathbf{z}$, где $\mathbf{z} \in \text{Ker } \mathcal{A} = (\text{Im } \mathcal{A}^*)^\perp$.

Теорема (нормальное псевдорешение)

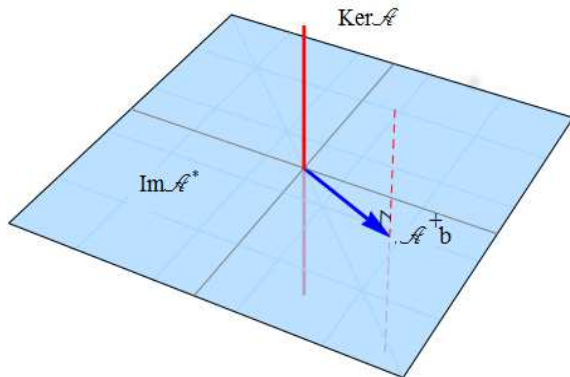
Для любой системы линейных уравнений $Ax = \mathbf{b}$ над полем $F \in \{\mathbb{R}, \mathbb{C}\}$ формула $\mathbf{x} = A^+\mathbf{b}$ возвращает ее нормальное псевдорешение.

Доказательство. Пусть A — $n \times k$ -матрица. Введем линейное отображение $\mathcal{A}: F^k \rightarrow F^n$, определяемое матрицей A по правилу $\mathcal{A}\mathbf{u} := A\mathbf{u}$; тогда матрица A^+ отвечает отображению \mathcal{A}^+ . По построению $A^+ = A_0^{-1}P_0$, где A_0 и P_0 — матрицы обратимого отображения \mathcal{A}_0 и ортопроектора \mathcal{P}_0 соответственно. $P_0\mathbf{b}$ — это ортогональная проекция столбца \mathbf{b} на подпространство $\text{Im } \mathcal{A}$, поэтому псевдорешения системы $Ax = \mathbf{b}$ — это столбцы $\mathbf{x} \in F^k$, удовлетворяющие $Ax = P_0\mathbf{b}$. Столбец $A^+\mathbf{b} \in \text{Im } \mathcal{A}^*$ этому равенству удовлетворяет, так как $A(A^+\mathbf{b}) = A_0(A_0^{-1}P_0\mathbf{b}) = P_0\mathbf{b}$. Итак, $A^+\mathbf{b}$ — псевдорешение, и любое псевдорешение представимо в виде $\mathbf{x} = A^+\mathbf{b} + \mathbf{z}$, где $\mathbf{z} \in \text{Ker } \mathcal{A} = (\text{Im } \mathcal{A}^*)^\perp$. По теореме Пифагора $|\mathbf{x}|^2 = |A^+\mathbf{b}|^2 + |\mathbf{z}|^2$, откуда длина \mathbf{x} минимальна при $\mathbf{z} = \mathbf{0}$, т.е. $\mathbf{x} = A^+\mathbf{b}$. □

Вычислять псевдообратную матрицу нелегко, но доказательство теоремы о нормальном псевдорешении подсказывает практически пригодный метод для нахождения такого псевдорешения.

Практический алгоритм для нормального псевдорешения

Вычислять псевдообратную матрицу нелегко, но доказательство теоремы о нормальном псевдорешении подсказывает практически пригодный метод для нахождения такого псевдорешения.



Среди всех псевдорешений (красная пунктирная прямая на рисунке) нормальное выделяется тем, что оно ортогонально ядру A , т.е. пространству решений однородной системы $Ax = 0$.

Вычислять псевдообратную матрицу нелегко, но доказательство теоремы о нормальном псевдорешении подсказывает практически пригодный метод для нахождения такого псевдорешения.

Среди всех псевдорешений нормальное выделяется тем, что оно ортогонально ядру \mathcal{A} , т.е. пространству решений однородной системы $Ax = 0$. Найдем базис этого ядра, т.е. фундаментальную систему решений, и составим из столбцов фундаментальной системы матрицу F .

Вычислять псевдообратную матрицу нелегко, но доказательство теоремы о нормальном псевдорешении подсказывает практически пригодный метод для нахождения такого псевдорешения.

Среди всех псевдорешений нормальное выделяется тем, что оно ортогонально ядру \mathcal{A} , т.е. пространству решений однородной системы $A\mathbf{x} = \mathbf{0}$. Найдем базис этого ядра, т.е. фундаментальную систему решений, и составим из столбцов фундаментальной системы матрицу F . Нормальное псевдорешение системы $A\mathbf{x} = \mathbf{b}$ будет единственным одновременным решением объединенной системы
$$\begin{cases} A^*A\mathbf{x} = A^*\mathbf{b}, \\ F^*\mathbf{x} = \mathbf{0}. \end{cases}$$

Теорема (Пенроуз)

Отображение \mathcal{A}^+ удовлетворяет следующим четырем тождествам:

$$(\mathcal{A}\mathcal{A}^+)^* = \mathcal{A}\mathcal{A}^+ \quad (1),$$

$$\mathcal{A}\mathcal{A}^+\mathcal{A} = \mathcal{A} \quad (3),$$

$$(\mathcal{A}^+\mathcal{A})^* = \mathcal{A}^+\mathcal{A} \quad (2),$$

$$\mathcal{A}^+\mathcal{A}\mathcal{A}^+ = \mathcal{A}^+ \quad (4).$$

Теорема (Пенроуз)

Отображение \mathcal{A}^+ удовлетворяет следующим четырем тождествам:

$$(\mathcal{A}\mathcal{A}^+)^* = \mathcal{A}\mathcal{A}^+ \quad (1),$$

$$\mathcal{A}\mathcal{A}^+\mathcal{A} = \mathcal{A} \quad (3),$$

$$(\mathcal{A}^+\mathcal{A})^* = \mathcal{A}^+\mathcal{A} \quad (2),$$

$$\mathcal{A}^+\mathcal{A}\mathcal{A}^+ = \mathcal{A}^+ \quad (4).$$

Теорема (Пенроуз)

Отображение \mathcal{A}^+ удовлетворяет следующим четырем тождествам:

$$(\mathcal{A}\mathcal{A}^+)^* = \mathcal{A}\mathcal{A}^+ \quad (1),$$

$$\mathcal{A}\mathcal{A}^+\mathcal{A} = \mathcal{A} \quad (3),$$

$$(\mathcal{A}^+\mathcal{A})^* = \mathcal{A}^+\mathcal{A} \quad (2),$$

$$\mathcal{A}^+\mathcal{A}\mathcal{A}^+ = \mathcal{A}^+ \quad (4).$$

Обратно, если некоторое линейное отображение \mathcal{B} удовлетворяет тождествам (1)–(4) с \mathcal{B} , подставленным вместо \mathcal{A}^+ , то $\mathcal{B} = \mathcal{A}^+$.

Теорема (Пенроуз)

Отображение \mathcal{A}^+ удовлетворяет следующим четырем тождествам:

$$(\mathcal{A}\mathcal{A}^+)^* = \mathcal{A}\mathcal{A}^+ \quad (1), \quad \mathcal{A}\mathcal{A}^+\mathcal{A} = \mathcal{A} \quad (3),$$

$$(\mathcal{A}^+\mathcal{A})^* = \mathcal{A}^+\mathcal{A} \quad (2), \quad \mathcal{A}^+\mathcal{A}\mathcal{A}^+ = \mathcal{A}^+ \quad (4).$$

Обратно, если некоторое линейное отображение \mathcal{B} удовлетворяет тождествам (1)–(4) с \mathcal{B} , подставленным вместо \mathcal{A}^+ , то $\mathcal{B} = \mathcal{A}^+$.

Доказательство. Тождества (1)–(4) легко выводятся из предложенной конструкции для \mathcal{A}^+ .

Теорема (Пенроуз)

Отображение \mathcal{A}^+ удовлетворяет следующим четырем тождествам:

$$(\mathcal{A}\mathcal{A}^+)^* = \mathcal{A}\mathcal{A}^+ \quad (1),$$

$$\mathcal{A}\mathcal{A}^+\mathcal{A} = \mathcal{A} \quad (3),$$

$$(\mathcal{A}^+\mathcal{A})^* = \mathcal{A}^+\mathcal{A} \quad (2),$$

$$\mathcal{A}^+\mathcal{A}\mathcal{A}^+ = \mathcal{A}^+ \quad (4).$$

Обратно, если некоторое линейное отображение \mathcal{B} удовлетворяет тождествам (1)–(4) с \mathcal{B} , подставленным вместо \mathcal{A}^+ , то $\mathcal{B} = \mathcal{A}^+$.

Доказательство. Тождества (1)–(4) легко выводятся из предложенной конструкции для \mathcal{A}^+ . Например, легко подсчитать, что $\mathcal{A}\mathcal{A}^+ = \mathcal{P}_0$, а ортопроектор \mathcal{P}_0 самосопряжен.

Теорема (Пенроуз)

Отображение \mathcal{A}^+ удовлетворяет следующим четырем тождествам:

$$(\mathcal{A}\mathcal{A}^+)^* = \mathcal{A}\mathcal{A}^+ \quad (1), \quad \mathcal{A}\mathcal{A}^+\mathcal{A} = \mathcal{A} \quad (3),$$

$$(\mathcal{A}^+\mathcal{A})^* = \mathcal{A}^+\mathcal{A} \quad (2), \quad \mathcal{A}^+\mathcal{A}\mathcal{A}^+ = \mathcal{A}^+ \quad (4).$$

Обратно, если некоторое линейное отображение \mathcal{B} удовлетворяет тождествам (1)–(4) с \mathcal{B} , подставленным вместо \mathcal{A}^+ , то $\mathcal{B} = \mathcal{A}^+$.

Допустим, что \mathcal{B}_1 и \mathcal{B}_2 удовлетворяет тождествам (1)–(4).

Теорема (Пенроуз)

Отображение \mathcal{A}^+ удовлетворяет следующим четырем тождествам:

$$(\mathcal{A}\mathcal{A}^+)^* = \mathcal{A}\mathcal{A}^+ \quad (1), \quad \mathcal{A}\mathcal{A}^+\mathcal{A} = \mathcal{A} \quad (3),$$

$$(\mathcal{A}^+\mathcal{A})^* = \mathcal{A}^+\mathcal{A} \quad (2), \quad \mathcal{A}^+\mathcal{A}\mathcal{A}^+ = \mathcal{A}^+ \quad (4).$$

Обратно, если некоторое линейное отображение \mathcal{B} удовлетворяет тождествам (1)–(4) с \mathcal{B} , подставленным вместо \mathcal{A}^+ , то $\mathcal{B} = \mathcal{A}^+$.

Допустим, что \mathcal{B}_1 и \mathcal{B}_2 удовлетворяет тождествам (1)–(4). Тогда

$$\begin{aligned} \mathcal{A}\mathcal{B}_1 &\stackrel{(3)}{=} (\mathcal{A}\mathcal{B}_2\mathcal{A})\mathcal{B}_1 = (\mathcal{A}\mathcal{B}_2)(\mathcal{A}\mathcal{B}_1) \stackrel{(1)}{=} (\mathcal{A}\mathcal{B}_2)^*(\mathcal{A}\mathcal{B}_1)^* = \\ &= \mathcal{B}_2^*\mathcal{A}^*(\mathcal{A}\mathcal{B}_1)^* = \mathcal{B}_2^*(\mathcal{A}\mathcal{B}_1\mathcal{A})^* \stackrel{(3)}{=} \mathcal{B}_2^*\mathcal{A}^* = (\mathcal{A}\mathcal{B}_2)^* \stackrel{(1)}{=} \mathcal{A}\mathcal{B}_2. \end{aligned}$$

Теорема (Пенроуз)

Отображение \mathcal{A}^+ удовлетворяет следующим четырем тождествам:

$$(\mathcal{A}\mathcal{A}^+)^* = \mathcal{A}\mathcal{A}^+ \quad (1), \quad \mathcal{A}\mathcal{A}^+\mathcal{A} = \mathcal{A} \quad (3),$$

$$(\mathcal{A}^+\mathcal{A})^* = \mathcal{A}^+\mathcal{A} \quad (2), \quad \mathcal{A}^+\mathcal{A}\mathcal{A}^+ = \mathcal{A}^+ \quad (4).$$

Обратно, если некоторое линейное отображение \mathcal{B} удовлетворяет тождествам (1)–(4) с \mathcal{B} , подставленным вместо \mathcal{A}^+ , то $\mathcal{B} = \mathcal{A}^+$.

Допустим, что \mathcal{B}_1 и \mathcal{B}_2 удовлетворяет тождествам (1)–(4). Тогда

$$\begin{aligned} \mathcal{A}\mathcal{B}_1 &\stackrel{(3)}{=} (\mathcal{A}\mathcal{B}_2\mathcal{A})\mathcal{B}_1 = (\mathcal{A}\mathcal{B}_2)(\mathcal{A}\mathcal{B}_1) \stackrel{(1)}{=} (\mathcal{A}\mathcal{B}_2)^*(\mathcal{A}\mathcal{B}_1)^* = \\ &= \mathcal{B}_2^*\mathcal{A}^*(\mathcal{A}\mathcal{B}_1)^* = \mathcal{B}_2^*(\mathcal{A}\mathcal{B}_1\mathcal{A})^* \stackrel{(3)}{=} \mathcal{B}_2^*\mathcal{A}^* = (\mathcal{A}\mathcal{B}_2)^* \stackrel{(1)}{=} \mathcal{A}\mathcal{B}_2. \end{aligned}$$

Аналогично проверяется, что $\mathcal{B}_1\mathcal{A} = \mathcal{B}_2\mathcal{A}$.

Теорема (Пенроуз)

Отображение \mathcal{A}^+ удовлетворяет следующим четырем тождествам:

$$(\mathcal{A}\mathcal{A}^+)^* = \mathcal{A}\mathcal{A}^+ \quad (1), \quad \mathcal{A}\mathcal{A}^+\mathcal{A} = \mathcal{A} \quad (3),$$

$$(\mathcal{A}^+\mathcal{A})^* = \mathcal{A}^+\mathcal{A} \quad (2), \quad \mathcal{A}^+\mathcal{A}\mathcal{A}^+ = \mathcal{A}^+ \quad (4).$$

Обратно, если некоторое линейное отображение \mathcal{B} удовлетворяет тождествам (1)–(4) с \mathcal{B} , подставленным вместо \mathcal{A}^+ , то $\mathcal{B} = \mathcal{A}^+$.

Допустим, что \mathcal{B}_1 и \mathcal{B}_2 удовлетворяет тождествам (1)–(4). Тогда

$$\begin{aligned} \mathcal{A}\mathcal{B}_1 &\stackrel{(3)}{=} (\mathcal{A}\mathcal{B}_2\mathcal{A})\mathcal{B}_1 = (\mathcal{A}\mathcal{B}_2)(\mathcal{A}\mathcal{B}_1) \stackrel{(1)}{=} (\mathcal{A}\mathcal{B}_2)^*(\mathcal{A}\mathcal{B}_1)^* = \\ &= \mathcal{B}_2^*\mathcal{A}^*(\mathcal{A}\mathcal{B}_1)^* = \mathcal{B}_2^*(\mathcal{A}\mathcal{B}_1\mathcal{A})^* \stackrel{(3)}{=} \mathcal{B}_2^*\mathcal{A}^* = (\mathcal{A}\mathcal{B}_2)^* \stackrel{(1)}{=} \mathcal{A}\mathcal{B}_2. \end{aligned}$$

Аналогично проверяется, что $\mathcal{B}_1\mathcal{A} = \mathcal{B}_2\mathcal{A}$. Теперь имеем

$$\mathcal{B}_1 \stackrel{(4)}{=} \mathcal{B}_1\mathcal{A}\mathcal{B}_1 = \mathcal{B}_1\mathcal{A}\mathcal{B}_2 = \mathcal{B}_2\mathcal{A}\mathcal{B}_2 \stackrel{(4)}{=} \mathcal{B}_2. \quad \square$$

Зная SVD матрицы, легко построить псевдообратную к ней.

Зная SVD матрицы, легко построить псевдообратную к ней.
Для простоты формулировок ограничимся евклидовым случаем.

Зная SVD матрицы, легко построить псевдообратную к ней.
Для простоты формулировок ограничимся евклидовым случаем.
(В унитарном случае нужно только поменять T на $*$.)

Зная SVD матрицы, легко построить псевдообратную к ней.
Для простоты формулировок ограничимся евклидовым случаем.
(В унитарном случае нужно только поменять T на $*$.)

Предложение (псевдообратная матрица через SVD)

Пусть M — $k \times n$ -матрица ранга r над полем \mathbb{R} , а $M = RAS$ — ее SVD, где A — $k \times n$ -матрица вида (#) с ненулевыми числами $\mu_1, \mu_2, \dots, \mu_r$ на диагонали. Тогда $M^+ = S^T A^+ R^T$, где A^+ — $n \times k$ -матрица вида (#) с ненулевыми числами $\mu_1^{-1}, \mu_2^{-1}, \dots, \mu_r^{-1}$ на диагонали.

Итак, чтобы найти псевдообратную матрицу по SVD данной матрицы, надо транспонировать SVD и обратить сингулярные числа.

Итак, чтобы найти псевдообратную матрицу по SVD данной матрицы, надо транспонировать SVD и обратить сингулярные числа.

$$M = R \begin{pmatrix} \mu_1 & 0 & \dots & 0 & 0 & \dots \\ 0 & \mu_2 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \mu_r & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots \end{pmatrix} S$$

Итак, чтобы найти псевдообратную матрицу по SVD данной матрицы, надо транспонировать SVD и обратить сингулярные числа.

$$M = R \begin{pmatrix} \mu_1 & 0 & \dots & 0 & 0 & \dots \\ 0 & \mu_2 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \mu_r & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots \end{pmatrix} S \implies M^+ = S^T \begin{pmatrix} \mu_1^{-1} & 0 & \dots & 0 & 0 & \dots \\ 0 & \mu_2^{-1} & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \mu_r^{-1} & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots \end{pmatrix} R^T$$

Итак, чтобы найти псевдообратную матрицу по SVD данной матрицы, надо транспонировать SVD и обратить сингулярные числа.

$$M = R \begin{pmatrix} \mu_1 & 0 & \dots & 0 & 0 & \dots \\ 0 & \mu_2 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \mu_r & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots \end{pmatrix} S \implies M^+ = S^T \begin{pmatrix} \mu_1^{-1} & 0 & \dots & 0 & 0 & \dots \\ 0 & \mu_2^{-1} & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \mu_r^{-1} & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots \end{pmatrix} R^T$$

Чтобы убедиться, что формула $M^+ = S^T A^+ R^T$ действительно возвращает псевдообратную матрицу, достаточно подставить выражение $S^T A^+ R^T$ в тождества (1)–(4) и проверить, что они выполняются.

Итак, чтобы найти псевдообратную матрицу по SVD данной матрицы, надо транспонировать SVD и обратить сингулярные числа.

$$M = R \begin{pmatrix} \mu_1 & 0 & \dots & 0 & 0 & \dots \\ 0 & \mu_2 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \mu_r & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots \end{pmatrix} S \implies M^+ = S^T \begin{pmatrix} \mu_1^{-1} & 0 & \dots & 0 & 0 & \dots \\ 0 & \mu_2^{-1} & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \mu_r^{-1} & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots \end{pmatrix} R^T$$

Чтобы убедиться, что формула $M^+ = S^T A^+ R^T$ действительно возвращает псевдообратную матрицу, достаточно подставить выражение $S^T A^+ R^T$ в тождества (1)–(4) и проверить, что они выполняются. Например, проверим равенство $MM^+M = M$:

$$\underbrace{RAS}_M \underbrace{S^T A^+ R^T}_{\text{кандидат в } M^+} \underbrace{RAS}_M$$

Итак, чтобы найти псевдообратную матрицу по SVD данной матрицы, надо транспонировать SVD и обратить сингулярные числа.

$$M = R \begin{pmatrix} \mu_1 & 0 & \dots & 0 & 0 & \dots \\ 0 & \mu_2 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \mu_r & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots \end{pmatrix} S \implies M^+ = S^T \begin{pmatrix} \mu_1^{-1} & 0 & \dots & 0 & 0 & \dots \\ 0 & \mu_2^{-1} & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \mu_r^{-1} & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots \end{pmatrix} R^T$$

Чтобы убедиться, что формула $M^+ = S^T A^+ R^T$ действительно возвращает псевдообратную матрицу, достаточно подставить выражение $S^T A^+ R^T$ в тождества (1)–(4) и проверить, что они выполняются. Например, проверим равенство $MM^+M = M$:

$$\underbrace{RAS}_M \underbrace{S^T A^+ R^T}_{\text{кандидат в } M^+} \underbrace{RAS}_M = RA \overbrace{SS^T}^E A^+ \overbrace{R^T R}^E AS = RAA^+AS$$

Итак, чтобы найти псевдообратную матрицу по SVD данной матрицы, надо транспонировать SVD и обратить сингулярные числа.

$$M = R \begin{pmatrix} \mu_1 & 0 & \dots & 0 & 0 & \dots \\ 0 & \mu_2 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \mu_r & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots \end{pmatrix} S \implies M^+ = S^T \begin{pmatrix} \mu_1^{-1} & 0 & \dots & 0 & 0 & \dots \\ 0 & \mu_2^{-1} & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \mu_r^{-1} & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \dots & \vdots & \vdots & \ddots \end{pmatrix} R^T$$

Чтобы убедиться, что формула $M^+ = S^T A^+ R^T$ действительно возвращает псевдообратную матрицу, достаточно подставить выражение $S^T A^+ R^T$ в тождества (1)–(4) и проверить, что они выполняются. Например, проверим равенство $MM^+M = M$:

$$\underbrace{RAS}_M \underbrace{S^T A^+ R^T}_{\text{кандидат в } M^+} \underbrace{RAS}_M = RA \overbrace{SS^T}^E A^+ \overbrace{R^T R}^E AS = RAA^+AS = RAS = M.$$

В многих практических задачах (сжатие данных, обработка сигналов, метод главных компонент, латентно-семантическое индексирование и т.д.) нужно приблизить матрицу M некоторой другой матрицей M_d с заранее заданным рангом d .

В многих практических задачах (сжатие данных, обработка сигналов, метод главных компонент, латентно-семантическое индексирование и т.д.) нужно приблизить матрицу M некоторой другой матрицей M_d с заранее заданным рангом d . При этом стремятся минимизировать $\|M - M_d\|^2$.

В многих практических задачах (сжатие данных, обработка сигналов, метод главных компонент, латентно-семантическое индексирование и т.д.) нужно приблизить матрицу M некоторой другой матрицей M_d с заранее заданным рангом d . При этом стремятся минимизировать $\|M - M_d\|^2$. (Под длиной $\|A\|$ матрицы A здесь понимается длина ее **векторизации**, т.е. длина вектора, который получится если вытянуть матрицы в строку.)

В многих практических задачах (сжатие данных, обработка сигналов, метод главных компонент, латентно-семантическое индексирование и т.д.) нужно приблизить матрицу M некоторой другой матрицей M_d с заранее заданным рангом d . При этом стремятся минимизировать $\|M - M_d\|^2$. (Под длиной $\|A\|$ матрицы A здесь понимается длина ее **векторизации**, т.е. длина вектора, который получится если вытянуть матрицы в строку.)

Теорема (Карл Эккарт, Гейл Янг, 1936)

Пусть M — $k \times n$ -матрица ранга r над полем \mathbb{R} , а $M = RAS$ — ее SVD, где A — $k \times n$ -матрица вида (#) с числами $\mu_1 \geq \mu_2 \geq \dots \geq \mu_r$ на диагонали, и пусть $d < r$. Тогда матрица M_d ранга d с наименьшей возможной величиной $\|M - M_d\|^2$ получается как $M_d := R_d A_d S_d$, где A_d — диагональная $d \times d$ -матрица с числами $\mu_1 \geq \mu_2 \geq \dots \geq \mu_d$ на диагонали, R_d — $k \times d$ -матрица, образованная первыми d столбцами матрицы R , а S_d — $d \times n$ -матрица, образованная первыми d строками матрицы S .

В многих практических задачах (сжатие данных, обработка сигналов, метод главных компонент, латентно-семантическое индексирование и т.д.) нужно приблизить матрицу M некоторой другой матрицей M_d с заранее заданным рангом d . При этом стремятся минимизировать $\|M - M_d\|^2$. (Под длиной $\|A\|$ матрицы A здесь понимается длина ее **векторизации**, т.е. длина вектора, который получится если вытянуть матрицы в строку.)

Теорема (Карл Эккарт, Гейл Янг, 1936)

Пусть M — $k \times n$ -матрица ранга r над полем \mathbb{R} , а $M = RAS$ — ее SVD, где A — $k \times n$ -матрица вида (#) с числами $\mu_1 \geq \mu_2 \geq \dots \geq \mu_r$ на диагонали, и пусть $d < r$. Тогда матрица M_d ранга d с наименьшей возможной величиной $\|M - M_d\|^2$ получается как $M_d := R_d A_d S_d$, где A_d — диагональная $d \times d$ -матрица с числами $\mu_1 \geq \mu_2 \geq \dots \geq \mu_d$ на диагонали, R_d — $k \times d$ -матрица, образованная первыми d столбцами матрицы R , а S_d — $d \times n$ -матрица, образованная первыми d строками матрицы S .

Итак, нужно оставить d первых сингулярных чисел, а остальные занулить.