Математика для Data Science. Линейная алгебра.
Решения задач

Содержание

Замечание. Вот этим цветом отмечены ссылки на страницы внутри этого файла.

Комплексные числа

Задача 1

Неформально. Рассмотрим конкретное комплексное число z_1. Мы смотрим на все комплексные числа как на векторы из $\mathbb{R}^2$. Будем доказывать, что умножение на z_1 делает следующее:

- растягивает все векторы в $|z_1|$ раз
- и поворачивает все векторы на угол $\arg(z)$ против часовой стрелки.

Ясно, что утверждение выше это то же самое, что теорема с предыдущего шага (только по-другому сформулированная).

А теперь формально.

1. Дано фиксированное комплексное число $z_1 = a + bi$. Мы рассматриваем отображение $m_{z_1} : \mathbb{C} \to \mathbb{C}$, которое отправляет каждое число z в произведение $z_1 \cdot z$. Вспомним, что $\mathbb{C}$ можно воспринимать как $\mathbb{R}^2$. Поэтому и наше отображение можно воспринимать как $m_{z_1} : \mathbb{R}^2 \to \mathbb{R}^2$. Вычислите, куда это отображение отправляет произвольное комплексное число $z = c + di = (c, d) \in \mathbb{R}^2$. Докажите, что отображение m_{z_1} из $\mathbb{R}^2$ в $\mathbb{R}^2$ является линейным и выпишите его матрицу в терминах a и b.

2. Выразите a и b через $r := |z|$ и $\alpha := \arg(z)$. Выпишите матрицу отображения m_{z_1} в терминах r и α.

3. Докажите теорему с предыдущего шага.

Подсказка. Вам может пригодиться шаг из урока про ортогональные матрицы.

Решение.

1. Пусть $z_1 = a + ib$, $z = c + di$. Тогда под действием отображения m_{z_1} число z переходит в $z_1 \cdot z = (a + ib)(c + di) = (ac - bd) + i(bc + ad)$. Если воспринимать $\mathbb{C}$ как $\mathbb{R}^2$, то $m(c, d) = (ac - bd, bc + ad)$. Итак, координаты отображения линейны по c и d, значит, само отображение тоже линейно.

   Матрица отображения равна $M = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. Действительно, $Mz = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} ac - bd \\ bc + ad \end{pmatrix}$

2. Согласно задаче прошлой недели $c = r\cos\alpha$ и $d = r\sin\alpha$.

   Пусть $R := |z_1|$, $\beta := \arg(z_1)$, то есть $\mathrm{Re}(z_1) = a = R\cos\beta$, $\mathrm{Im}(z_1) = b = R\sin\beta$. Тогда $\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} R\cos\beta & -R\sin\beta \\ R\sin\beta & R\cos\beta \end{pmatrix} \begin{pmatrix} r\cos\alpha \\ r\sin\alpha \end{pmatrix} = \begin{pmatrix} rR\cos\alpha\cos\beta - rR\cos\alpha\sin\beta \\ rR\cos\alpha\sin\beta + rR\sin\alpha\cos\beta \end{pmatrix} = \begin{pmatrix} rR\cos(\alpha+\beta) \\ rR\sin(\alpha+\beta) \end{pmatrix}$

3. Действительно, в предыдущем пункте видно, что при умножении комплексных чисел их модули перемножаются, а аргументы складываются.

Задача 2

1. Найдите число z^k, зная модуль и аргумент z.

2. Найдите какое-нибудь число z, такое что $z^6 = -100$.

3. Докажите, что у любого ненулевого комплексного числа есть обратное. То есть для любого ненулевого $z \in \mathbb{C}$ найдётся число z^{-1}, такое что $z \cdot z^{-1} = 1$.

В целом, мы доказали, что $\mathbb{C}$ это *поле*. То есть, в $\mathbb{C}$ есть 0 и 1, а также в $\mathbb{C}$ можно складывать, вычитать, умножать и делить.

Подсказка. В этих задачах поможет выражать комплексные числа через их модуль и аргумент.

Решение.

1. Пусть $z = r\cos\alpha + ir\sin\alpha$. Тогда по теореме из предыдущей задачи $z^2 = z \cdot z = r^2\cos(2\alpha) + ir^2\sin(2\alpha)$ (модули перемножились, а аргументы сложились). Отсюда следует, что $z^3 = r^3\cos(3\alpha) + ir^3\sin(3\alpha)$. Продолжая так далее, получим $z^k = r^k\cos(k\alpha) + ir^k\sin(k\alpha)$. То есть $|z^k| = |z|^k$ и $\arg(z^k) = k \cdot \arg(z)$.

2. Ищем z в виде $z = r\cos\alpha + ir\sin\alpha$, тогда по предыдущей задаче $z^6 = r^6(\cos(6\alpha) + i\sin(6\alpha))$. Запишем -100 как комплексное число с модулем 100 и аргументом $-\pi$: а именно, $-100 = 100(\cos\pi + i\sin\pi)$.

   Итак, $r^6(\cos(6\alpha) + i\sin(6\alpha)) = 100(\cos\pi + i\sin\pi)$.

   Значит, $r^6 = 100$, то есть $r = \sqrt[6]{100} = \sqrt[3]{10}$.

   Кроме того, $6\alpha = \pi + 2\pi k$, где $k \in Z$. Рассмотрим, например, случай $k = 1$. Тогда $\alpha = \frac{\pi}{2}$.

   Итак, мы получили $z = \sqrt[3]{10}(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}) = i\sqrt[3]{10}$.

   Замечание. При рассмотрении других $k \in \mathbb{Z}$ получились бы $z = -i\sqrt[3]{10}$, $z = \sqrt[3]{10}\left(\pm\frac{\sqrt{3}}{2} \pm \frac{i}{2}\right)$.

3. Пусть $z = r\cos\alpha + ir\sin\alpha$, найдём число z^{-1} в виде $z^{-1} = R\cos\beta + iR\sin\beta$.

   Тогда $z \cdot z^{-1} = rR(\cos(\alpha+\beta) + i\sin(\alpha+\beta)) = 1(\cos 0 + i\sin 0)$.

   Значит, $rR = 1$ и $\alpha + \beta = 0$. Другими словами, $R = r^{-1}$ и $\beta = -\alpha$. То есть $z^{-1} = r^{-1}(\cos(-\alpha) + i\sin(-\alpha)) = r^{-1}(\cos\alpha - i\sin\alpha)$

Собственные векторы

Задача 1

Докажите, что если $\vec{v}$ это собственный вектор A, то и $c\vec{v}$ это собственный вектор A, где c – любое ненулевое число.

Решение. Если $\vec{v}$ это собственный вектор A, то по определению $\exists\lambda : Av = \lambda v$. Тогда если c – любое ненулевое число, то $A(c\vec{v}) = cA\vec{v} = \vec{0}$, то есть $c\vec{v}$ — тоже собственный вектор.

Задача 2

Найдите собственные числа и собственные векторы матрицы $B = \begin{pmatrix} 2 & 8 \\ 3 & 4 \end{pmatrix}$.

Решение. Сначала найдём все собственные значения матрицы B.

$0 = \det(B - \lambda E) = \det\left(\begin{pmatrix} 2 & 8 \\ 3 & 4 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\right) = \det\begin{pmatrix} 2-\lambda & 8 \\ 3 & 4-\lambda \end{pmatrix} = (2-\lambda)(4-\lambda) - 24 = \lambda^2 - 6\lambda - 16$. Значит, $\lambda = -2$ или $\lambda = 8$.

Найдём теперь собственный вектор, соответствующий собственному числу -2. Это такой вектор $\vec{v}$, что $(B - (-2E))\vec{v} = 0$.

$(B + 2E) = \begin{pmatrix} 2+2 & 8 \\ 3 & 4+2 \end{pmatrix} = \begin{pmatrix} 4 & 8 \\ 3 & 6 \end{pmatrix}$. Собственный вектор равен $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$.

Теперь найдём собственный вектор, соответствующий $\lambda = 8$. $(B - 8E) = \begin{pmatrix} 2-8 & 8 \\ 3 & 4-8 \end{pmatrix} = \begin{pmatrix} -6 & 8 \\ 3 & -4 \end{pmatrix}$. Собственный вектор равен $\begin{pmatrix} 4 \\ 3 \end{pmatrix}$.

Задача 3

1. Докажите, что у матрицы $\begin{pmatrix} 1 & -10 \\ 10 & 1 \end{pmatrix}$ нет собственных чисел и собственных векторов (над $\mathbb{R}$)

2. Докажите, что при $0 < \alpha < \pi$ у матрицы $\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$ нет собственных чисел и собственных векторов (над $\mathbb{R}$). Что это значит геометрически?

Решение.

1. Найдём собственные числа так же, как ранее: $\det\left(\begin{pmatrix}1-\lambda & -10\\ 10 & 1-\lambda\end{pmatrix}\right) = (1-\lambda)^2+100 = \lambda^2-2\lambda+101$. Поскольку дискриминант отрицателен ($4-4\cdot 101<0$), корней над $\mathbb{R}$ у уравнения нет.

2. $\det\left(\begin{pmatrix}\cos(\alpha)-\lambda & -\sin(\alpha)\\ \sin(\alpha) & \cos(\alpha)-\lambda\end{pmatrix}\right) = (\cos(\alpha)-\lambda)^2+\sin^2\alpha = \cos^2\alpha+\sin^2\alpha+\lambda^2-2\lambda\cos\alpha = \lambda^2-2\lambda\cos\alpha+1$. Дискриминант равен $4(\cos^2\alpha-1)$. Если $0<\alpha<\pi$, то $\cos^2\alpha<1$ и дискриминант отрицателен, следовательно, собственных чисел и векторов не существует. Геометрически это означает, что при повороте на угол α нет неподвижных векторов.

Задача 4

Докажите, что каждому собственному числу соответствует хотя бы один собственный вектор. То есть если для числа λ выполнено $\det(A-\lambda E)=0$, то найдётся $\vec{v}$, такой что $A\vec{v}=\lambda\vec{v}$

Решение. Пусть для числа λ выполнено $\det(A-\lambda E)=0$. Тогда у матрицы $(A-\lambda E)$ неполный ранг. Значит, столбцы этой матрицы линейно зависимы и из коэффициентов этой линейной зависимости можно составить ненулевой вектор $\vec{v}$ такой, что $(A-\lambda E)\vec{v}=\vec{0}$. Это равносильно $A\vec{v}-(\lambda E)\vec{v}=\vec{0} \Leftrightarrow A\vec{v}-\lambda\vec{v}=\vec{0} \Leftrightarrow A\vec{v}=\lambda\vec{v}$.

Задача 5 (дополнительная)

Пусть $\vec{v}_1,\dots,\vec{v}_k$ это собственные векторы с различными собственными числами $\lambda_1,\dots,\lambda_k$. Докажите, что $\vec{v}_1,\dots,\vec{v}_k$ линейно независимы.

Решение. Пусть $\vec{v}_1,\dots,\vec{v}_k$ это собственные векторы матрицы A с различными собственными числами $\lambda_1,\dots,\lambda_k$. Рассмотрим сначала один вектор $\vec{v}_1$. Он не равен нулю, поэтому линейно независим.

Далее, допустим, что мы уже доказали, что векторы $\vec{v}_1,\dots,\vec{v}_{m-1}$ линейно независимы. Докажем от противного, что $\vec{v}_1,\dots,\vec{v}_m$ тоже линейно независимы. Действительно, пусть $\vec{v}_1,\dots,\vec{v}_m$ линейно зависимы. Тогда

$$\vec{v}_m = \alpha_1\vec{v}_1+\dots+\alpha_{m-1}\vec{v}_{m-1},$$

где хотя бы одно $\alpha_i\neq 0$.

Применим A к обеим частям равенства и, пользуясь тем, что $A\vec{v}_j=\lambda_j v_j$, получим:

$$\lambda_m\vec{v}_m = \alpha_1\lambda_1\vec{v}_1+\dots+\alpha_{m-1}\lambda_{m-1}\vec{v}_{m-1}$$

С другой стороны,

$$\lambda_m\vec{v}_m = \lambda_m(\alpha_1\vec{v}_1+\dots+\alpha_{m-1}\vec{v}_{m-1})$$

Вычитая предпоследнее равенство из последнего, получаем

$$\alpha_1(\lambda_m-\lambda_1)\vec{v}_1+\dots+\alpha_m(\lambda_m-\lambda_{m-1})\vec{v}_{m-1}=\vec{0}$$

Но $\vec{v}_1,\dots,\vec{v}_{m-1}$ линейно независимы, а значит все коэффициенты равны нулю: $\alpha_j(\lambda_m-\lambda_j)=0$, но так как хотя бы одно $\alpha_i\neq 0$, то $\lambda_m-\lambda_i=0$, что невозможно, так все все собственные числа по условию разные. Значит, наше предположение было неверно и $\vec{v}_1,\dots,\vec{v}_m$ линейно независимы.

Проведя приведённое выше рассуждение последовательно для $m=2$, затем $m=3$ и так далее до $m=k$, мы докажем требуемое.

Низкоранговое приближение матрицы

Задача 3

Давайте поймём, как взаимодействует ранг и операция умножения матриц. Во всех задачах ниже размеры матриц считайте любыми (но такими, что умножение определено).

1. Докажите, что ранг произведения двух матриц меньше или равен ранга каждой из них: $\mathrm{rank}(AB)\leq \min(\mathrm{rank}(A),\mathrm{rank}(B))$.

2. Докажите, что ранг произведения любого числа матриц меньше или равен ранга каждой из них: $\mathrm{rank}(A_1 \cdot A_2 \cdot \dots \cdot A_r) \le \min(\mathrm{rank}(A_1), \mathrm{rank}(A_2), \dots, \mathrm{rank}(A_r))$.

3. Приведите пример матриц A и B, для которых ранг произведения меньше ранга каждого из сомножителей: $\mathrm{rank}(AB) < \min(\mathrm{rank}(A), \mathrm{rank}(B))$.

Подсказка. Вам может помочь интерпретация ранга из задачи прошлых недель. Сколькими векторами порождён образ B? Сколькими векторами порождён образ A? Сколькими векторами порождён образ AB?

Решение.

1. Согласно задаче со второй недели, ранг матрицы A равен размерности образа A. То есть образ A порождён $\mathrm{rank}(A)$ векторами, а образ B порождён $\mathrm{rank}(B)$ векторами. Матрица AB соответствует композиции отображений A и B, поэтому образ AB не может быть порождён линейно независимыми векторами, которых больше, чем $\mathrm{rank}(A)$ или $\mathrm{rank}(B)$.

   Немного другими словами то же самое можно объяснить тем, что подпространство, порожденное векторами-столбцами матрицы AB, содержится в подпространстве, порожденном векторами-столбцами матрицы A и то же самое верно для матрицы B.

2. По первому пункту этой задачи $\mathrm{rank}(A_1 \cdot A_2 \cdot \dots \cdot A_r) \le \min(\mathrm{rank}(A_1), \mathrm{rank}(A_2 \cdot \dots \cdot A_r))$. Далее, опять же по первому пункту задачи $\mathrm{rank}(A_2 \cdot \dots \cdot A_r) \le \min(\mathrm{rank}(A_2), \mathrm{rank}(A_3 \cdot \dots \cdot A_r))$. Аналогично применяя первый пункт задачи и дальше, получим требуемое неравенство $\mathrm{rank}(A_1 \cdot A_2 \cdot \dots \cdot A_r) \le \min(\mathrm{rank}(A_1), \mathrm{rank}(A_2), \dots, \mathrm{rank}(A_r))$.

3. Например, можно рассмотреть $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ и $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Ясно, что $\mathrm{rank}(A) = \mathrm{rank}(B) = 1$. Но так как $AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, то $\mathrm{rank}(AB) = 0$.

Задача 4

Пусть дана матрица W размера n на m и ранга не больше 1. Тогда можно найти такие матрицы U и M, что $W = U^T M$. При этом U это матрица размера 1 на n, и M это матрица размера 1 на m. То есть U^T это столбец, а M это строка.

Рассмотрим два случая:

1. $\mathrm{rank} W = 0$. Тогда W это нулевая матрица и U^T это нулевой столбец высоты n и M это нулевая строка длины m.

2. $\mathrm{rank} W = 1$. Тогда стобцы матрицы W попарно линейно зависимы, а значит, матрица имеет вид $(\vec{u}, \lambda_2 \vec{u}_1, \dots, \lambda_n \vec{u}_1)$. То есть $W = \vec{u}_1(1, \lambda_2, \dots, \lambda_n)$ — искомое представление в виде произведения столбца на строку.

Задача 5

Пусть дана матрица W размера n на m и ранга не больше k. Тогда можно найти такие матрицы U и M, что $W = U^T M$. При этом U это матрица размера k на n, и M это матрица размера k на m.

Подсказка. Примените метод Гаусса.

Решение. Пусть ранг матрицы W размера n на m меньше или равен k. Тогда после алгоритма Гаусса в матрице останется k или меньше ненулевых столбцов. Составим из этих столбцов матрицу. При этом столбцы изначальной матрицы выражаются через получившиеся после метода Гаусса столбцы, и из соответствующих коэффициентов остаётся составить вторую матрицу.

Сингулярное разложение – SVD

Задача 1

Пусть A_1 и A_2 – обратимые матрицы.

1. Докажите, что $\text{rank}(A_1B) = \text{rank}(B)$ для любой B, если произведение A_1B определено.
2. Докажите, что $\text{rank}(BA_2) = \text{rank}(B)$ для любой B, если произведение BA_2 определено.

Как мы знаем, все ортогональные матрицы обратимы. В SVD нас будет интересовать случай, когда A_1 и A_2 это ортогональные матрицы.

Подсказка. Ранее мы доказали, что $\text{rank}(A_1B) \leq$ B.

Решение.

1. Ранее мы доказали, что $\text{rank}(A_1B) \leq \text{rank}(B)$.

   Верна следующая цепочка равенств: $B = EB = (A_1^{-1}A_1)B = A_1^{-1}(A_1B)$. Снова пользуясь неравенством для ранга произведения матриц, получаем, что $\text{rank}(B) \leq \text{rank}(A_1B)$.

   Итак, $\text{rank}(A_1B) \leq \text{rank}(B)$ и $\text{rank}(B) \leq \text{rank}(A_1B)$, значит, $\text{rank}(B) = \text{rank}(A_1B)$.

2. Этот пункт доказывается аналогично первому:

   $\text{rank}(BA_2) \leq \text{rank}(B)$

   С другой стороны, $B = BE = B(A_2A_2^{-1}) = (BA_2)A_2^{-1}$, а значит $\text{rank}(B) \leq \text{rank}(BA_2)$. А тогда $\text{rank}(BA_2) = \text{rank}(B)$

Задача 2

Давайте докажем, что умножение матрицы на ортогональную матрицу не меняет её норму Фробениуса.

Пусть X – произвольная матрица размера m на n. Пусть L и R – ортогональные матрицы размера m на m и n на n соответственно.

1. Докажите, что $||LX||_F^2 = ||X||_F^2$
2. Докажите, что $||XR||_F^2 = ||X||_F^2$

Эта задача позволит нам исключить из вычислений ортогональные матрицы U и V^T, сконцентрировавшись на понятной матрице Σ.

Подсказка. Ортогональное преобразование сохраняет длины.

Решение.

1. Пусть $\vec{x}_1, \ldots, \vec{x}_n$ — столбцы матрицы X. Тогда $||X||_F^2 = ||\vec{x}_1||^2 + \cdots + ||\vec{x}_n||^2$.

   Кроме того, $||LX||_F^2 = ||L\vec{x}_1||^2 + \cdots + ||L\vec{x}_n||^2$, и, продолжая цепочку равенств и пользуясь ортогональностью L, получаем $||LX||_F^2 = ||\vec{x}_1||^2 + \cdots + ||\vec{x}_n||^2 = ||X||_F^2$

2. Для любой матрицы A верно $||A^T||_F^2 = ||A||_F^2$, ведь в определении нормы просто поменяется порядок суммирования.

   Итак, $||XR||_F^2 = ||(XR)^T||_F^2 = ||R^TX^T||_F^2$. Если R ортогонально, то и R^T тоже ортогонально, ведь $R^TR = E$. Значит, по первому пункту этой задачи $||R^TX^T||_F^2 = ||X^T||_F^2 = ||X||_F^2$. Итого, $||XR||_F^2 = ||X||_F^2$.

Задача 3

Дана диагональная матрица Σ с $\sigma_1 \geq \ \sigma_2 \geq \sigma_3 \geq \dots \geq 0:$

$$\Sigma = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \ddots \end{pmatrix}$$

Мы ищем P такую что $||\Sigma - P||_F^2$ минимально, при ограничении $\operatorname{rank}(P) \leq k$. Докажите, что минимальное значение достигается при $P = \Sigma_k$ (матрицу Σ_k мы определили ранее на Степике).

Тем самым лучшее приближение матрицы $A = U\Sigma V^T$ это матрица $B = UPV^T = U\Sigma_k V^T =: A_k$.

Это завершает доказательство теоремы.

Подсказка. Может ли P быть недиагональной матрицей? Что должно стоять у P на диагонали?

Решение. Чтобы запись была короче, введём обозначение для диагональных матриц: $\Sigma := \operatorname{diag}\{\sigma_1, \dots, \sigma_l\}$. Поскольку в матрице Σ вне диагонали стоят нули, то и P мы будем искать тоже такого вида, ведь иначе к норме $||\Sigma - P||_F^2$ прибавятся слагаемые, соответствующие квадратам элементов вне диагонали. Итак, пусть $P = \operatorname{diag}\{p_1, \dots, p_l\}$, при этом так как $\operatorname{rk}(P) \leq k$, то ненулевых p_i не больше, чем k. Итак, мы хотим минимизировать $||\Sigma - P||_F^2 = \sum\limits_{i=1}^{l} (\sigma_i - p_i)^2$. Поскольку есть хотя бы $l - k$ нулевых p_i, то в этой сумме есть хотя бы $l - k$ слагаемых вида σ_i^2. Вспомним, что $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \sigma_l \geq 0$. Тогда в $||\Sigma - P||_F^2$ мы хотим оставить слагаемые, соответствующие наименьшим σ_i, а остальные занулить. Значит, $P = \operatorname{diag}\{\sigma_1, \dots, \sigma_k, 0 \dots, 0\}$, а это и есть Σ_k.

Задача 4

Давайте подумаем, какую часть информации мы потеряли, заменив A на A_k.

Наблюдения, которые понадобятся для решения этой задачи:

- $||A||_F^2 = ||U\Sigma V^T||_F^2 = ||\Sigma||_F^2$, так как умножение на ортогональные матрицы сохраняет норму (пятая устная задача).
- $||\Sigma||_F^2 = \sigma_1^2 + \sigma_2^2 + \dots$. Всего элементов на диагонали $\min(n, m)$, обозначим $l := \min(n, m)$. Тогда $||\Sigma||_F^2 = \sum\limits_{i=1}^{l} \sigma_i^2$.
- Тем самым $||A||_F^2 = \sum\limits_{i=1}^{l} \sigma_i^2$.

1) Докажите, что $||A_k||_F^2 = \sum\limits_{i=1}^{k} \sigma_i^2$.

2) Докажите, что $||A - A_k||_F^2 = \sum\limits_{i=k+1}^{l} \sigma_i^2$. То есть $||A||_F^2 = ||A - A_k||_F^2 + ||A_k||_F^2$.

Решение.

1) Из определения матрицы A_k следует, что $||A_k||_F^2 = ||U\Sigma_k V^T||_F^2$. Согласно пятой устной задаче этого урока, умножение на ортогональные матрицы не меняет норму Фробениуса, следовательно, $||U\Sigma_k V^T||_F^2 = ||\Sigma_k||_F^2 = \sum\limits_{i=1}^{k} \sigma_i^2$.

2) $||A - A_k||_F^2 = ||U\Sigma V^T - U\Sigma_k V^T||_F^2 = ||U(\Sigma - \Sigma_k)V^T||_F^2 = ||\Sigma - \Sigma_k||_F^2 = \sum\limits_{i=k+1}^{l} \sigma_i^2$. Здесь мы снова воспользовались тем, что умножение на ортогональные матрицы не меняет норму Фробениуса.