Харківський національний університет
імені В. Н. Каразіна

“ХЕМОІНФОРМАТИКА ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ
для хіміків”

Лекція №5
“Некоректні задачі матричної алгебри”

В. В. Іванов

Кафедра хімічного матеріалознавства

$$A\vec{X} = \vec{Y}$$

Однорідні $\vec{Y} = 0$

неоднорідні $\vec{Y} \neq 0$

- Якщо число рівнянь (m), **більше** числа змінних (n) – *перевизначена* система
- Якщо число рівнянь (m) , **менш** числа змінних (n) – *недовизначена* система
- Система *сумісна*, якщо у неї є хоча б одне рішення. У противному випадку вона *несумісна.*

$$m = n, \quad \det(A) \neq 0$$

$$\vec{X} = A^{-1}\vec{Y}$$

Приклад 1.

$$\begin{cases} 2x_1 + x_2 = 1 \\ x_1 + 2x_2 = 2 \end{cases} \qquad \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \qquad \det A = 4 - 1 = 3$$

$$A^{-1} = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \qquad \overset{F}{X} = A^{-1} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

розв'язок $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Вносимо невелике збурення $\begin{pmatrix} 1 \\ 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1.1 \\ 2 \end{pmatrix}$

$$\overset{F}{X} = A^{-1}\begin{pmatrix} 1.1 \\ 2 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 1.1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0.066666 \\ 0.966666 \end{pmatrix}$$

було таке $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ стало таке $\begin{pmatrix} 0.066666 \\ 0.966666 \end{pmatrix}$

Приклад 2

$$\begin{cases} x_1 + 10x_2 = 11 \\ 10x_1 + 101x_2 = 111 \end{cases} \qquad \begin{pmatrix} 1 & 10 \\ 10 & 101 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 11 \\ 111 \end{pmatrix}$$

$$A^{-1} = \begin{pmatrix} 101 & -10 \\ -10 & 1 \end{pmatrix}$$

рішення: $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$$A^{-1}\begin{pmatrix} 11.1 \\ 111 \end{pmatrix} = \begin{pmatrix} 101 & -10 \\ -10 & 1 \end{pmatrix}\begin{pmatrix} 11.1 \\ 111 \end{pmatrix} = \begin{pmatrix} 1121.1 - 1110 \\ -111 + 111 \end{pmatrix} = \begin{pmatrix} 11.1 \\ 0 \end{pmatrix}$$

було: $\boxed{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}$ стало: $\boxed{\begin{pmatrix} 11.1 \\ 0 \end{pmatrix}}$

$$A^{-1}\begin{pmatrix} 11 \\ 111.1 \end{pmatrix} = \begin{pmatrix} 101 & -10 \\ -10 & 1 \end{pmatrix}\begin{pmatrix} 11 \\ 111.1 \end{pmatrix} = \begin{pmatrix} 1111 - 1111 \\ -110 + 111.1 \end{pmatrix} = \boxed{\begin{pmatrix} 0 \\ 1.1 \end{pmatrix}}$$

Такі задачі звуться **погано обумовленими**

Число обумовленості матриці

$$Ax = Y \qquad A^{-1}Y = x$$

$$\| A \| \cdot \| x \| \geq \| Y \| \qquad \| A^{-1} \| \cdot \| Y \| \geq \| x \|$$

$$COND(A) = \frac{\|A\|}{\|A^{-1}\|}$$

$$A = \begin{pmatrix} 1 & 10 \\ 10 & 101 \end{pmatrix} \qquad \begin{vmatrix} 1-\lambda & 10 \\ 10 & 101-\lambda \end{vmatrix} = 0$$

$$(1-\lambda)(101-\lambda) - 10 \cdot 10 = 0$$

$$\lambda^2 - 102\lambda + 1 = 0$$

$$\lambda_{1,2} = \frac{102 \pm \sqrt{102^2 - 4 \cdot 1 \cdot 1}}{2} = \frac{102 \pm 101.9804}{2} = 51 \pm 50.9902$$

$$\lambda_{\max} = 101.9902, \quad \lambda_{\min} = 0.0098$$

$$\mathrm{COND}(A) = \frac{101.9902}{0.0098} = 1.04 \cdot 10^4$$

До першого прикладу

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

$$\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 0$$

$$(2-\lambda)^2 - 1 = 0$$

$$\lambda^2 - 4\lambda + 3 = 0$$

$$\lambda_{max} = 3, \quad \lambda_{min} = 1$$

$$\mathrm{COND(A)} = 3$$

Приклад 3

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \qquad \begin{cases} x_1 + x_2 = 5 \\ x_1 + x_2 = 5 \end{cases}$$

$\det(A) = 1 \cdot 1 - 1 \cdot 1 = 0.$

$$\begin{vmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix} = 0 \qquad \lambda^2 - 2\lambda = 0 \qquad \lambda_{max} = 2, \quad \lambda_{min} = 0$$

$$\mathrm{COND}(A) = \frac{2}{0} = \infty$$

$$x_1 + x_2 = 5 \qquad x_2 = 5 - x_1$$

$$x_1 = \alpha \qquad x_2 = 5 - \alpha \qquad \begin{pmatrix} \alpha \\ 5-\alpha \end{pmatrix}$$

Довільне рішення є точним !

$$Ax - y = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \alpha \\ 5-\alpha \end{pmatrix} - \begin{pmatrix} 5 \\ 5 \end{pmatrix} = \begin{pmatrix} \alpha + 5 - \alpha - 5 \\ \alpha + 5 - \alpha - 5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Знайдемо таке рішення, яке відповідає мінімальному x (за нормою) !

$$\begin{pmatrix} \alpha & 5-\alpha \end{pmatrix} \cdot \begin{pmatrix} \alpha \\ 5-\alpha \end{pmatrix} = \alpha^2 + (5-\alpha)^2 = 2\alpha^2 - 10\alpha + 25$$

$4\alpha - 10 = 0 \qquad \alpha = \frac{5}{2}$ **Тобто найкраще рішення:** $\frac{5}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2.5 \\ 2.5 \end{pmatrix}$

Можливий такий підхід

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \to \quad \widetilde{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1+\varepsilon \end{pmatrix}$$

$$\det(\widetilde{A}) = 1\cdot(1+\varepsilon) - 1\cdot 1 = \varepsilon$$

$$\widetilde{A}^{-1} = \frac{1}{\varepsilon}\begin{pmatrix} 1+\varepsilon & -1 \\ -1 & 1 \end{pmatrix}$$

$$X = A^{-1} Y = \frac{1}{\varepsilon}\begin{pmatrix} 1+\varepsilon & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \frac{1}{\varepsilon}\begin{pmatrix} (1+\varepsilon)y_1 - y_2 \\ -y_1 + y_2 \end{pmatrix}$$

$$\begin{cases} x_1 = \dfrac{(1+\varepsilon)y_1 - y_2}{\varepsilon} \\ x_2 = \dfrac{-y_1 + y_2}{\varepsilon} \end{cases}$$

Чому це погано ?

Коректність задачі $A\vec{X} = \vec{Y}$

Перша спроба визначення коректності математичного завдання належить французькому математику Жаку Адамару (1865-1963).

«Одна з фундаментальних проблем небесної механіки – проблема стійкості сонячної системи – потрапляє до категорії некоректних завдань………… кожну стійку траєкторію можна перетворити нескінченно малою зміною початкових даних у цілком нестійку траєкторію, що йде на нескінченність. В астрономічних задачах початкові дані відомі лише з певною помилкою. Але як не мала ця помилка, вона може спричинити повне і абсолютне збурення в результаті».

1901 рік.

Коректність за Адамаром

$$A\vec{X} = \vec{Y}$$

Завдання може вважатися коректним (або коректно поставленим) якщо:

•Рівність $AX_1 = AX_2$ для векторів X_1 і X_2 призводить до рівності $X_1 = X_2$ (умова єдності рішення);

•Зворотна матриця існує і однозначно визначена (умова сталості рішень);

•Рішення безперервно залежить від даних.

Решта завдань (некоректні) нефізичні.

Коректність за Тихоновим

$$A\vec{X} = \vec{Y}$$

•Апріорно відомо, що розв'язання задачі існує і належить апріорно заданій множині;

•Рішення єдине;

•Нескінченно малими варіаціями правої частини, які не виводять рішення із заданого класу рішень, відповідає нескінченно малі варіації рішення.

Стандартна схема рішення задачі - мінімізація функціоналу похибки

$$A\vec{X} = \vec{Y}$$

$$A\vec{X} - \vec{Y} = \vec{R}$$ Вектор **невязки**

$$\vec{R}\cdot\vec{R} = R^{+}\cdot R = (.....R......)\begin{pmatrix} 7 \\ R \\ 7 \end{pmatrix} = \| R \|$$

$$\mathbf{E(X)} = \| \mathbf{R} \| = \left\| \mathbf{AX - Y} \right\| = (\mathbf{X}^{+}\mathbf{A}^{+} - \mathbf{Y}^{+})(\mathbf{AX - Y})$$

$$\mathbf{E(X)} = (\mathbf{X}^{+}\mathbf{A}^{+} - \mathbf{Y}^{+})(\mathbf{AX - Y}) = \mathbf{X}^{+}\mathbf{A}^{+}\mathbf{AX} - \mathbf{X}^{+}\mathbf{A}^{+}\mathrm{Y} - \mathrm{Y}^{+}\mathbf{AX}$$

$$\frac{\partial}{\partial \mathbf{X}^{+}} \mathrm{E(X)} = A^{+}AX - A^{+}Y \qquad A^{+}AX - A^{+}Y = 0$$

$$A^{+}AX = A^{+}Y \quad \rightarrow \quad \boxed{X = (A^{+}A)^{-1}A^{+}Y}$$

Метод Регуляризації Тихонова

За Тихоновим:

$$F_\varepsilon(X) = \|AX - Y\| + \varepsilon\|X\| = (X^+A^+ - Y^+)(AX - Y) + \varepsilon X^+X$$

$$\frac{\partial}{\partial \mathbf{X}^+}\mathbf{F}_\varepsilon(\mathbf{X}) = \frac{\partial}{\partial \mathbf{X}^+}(\mathbf{X}^+\mathbf{A}^+\mathbf{A}\mathbf{X} - \mathbf{X}^+\mathbf{A}^+\mathbf{Y} - \mathbf{Y}^+\mathbf{A}\mathbf{X} + \mathbf{Y}^+\mathbf{Y} + \varepsilon\mathbf{X}^+\mathbf{X}) = \mathbf{0}$$

$$A^+AX - A^+Y + \varepsilon X = 0$$

$$(A^+A + \varepsilon I)X = A^+Y \longrightarrow \boxed{X = (A^+A + \varepsilon I)^{-1}A^+Y}$$

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \qquad \left(\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \varepsilon \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$\left(\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} + \begin{pmatrix} \varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix} \right) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 + y_2 \\ y_1 + y_2 \end{pmatrix} \qquad \begin{pmatrix} 2+\varepsilon & 2 \\ 2 & 2+\varepsilon \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 + y_2 \\ y_1 + y_2 \end{pmatrix}$$

$$\det(\widetilde{A}) = (2+\varepsilon)^2 - 4 = 4\varepsilon + \varepsilon^2 = \varepsilon(\varepsilon + 4) \qquad \widetilde{A}^{-1} = \frac{1}{\varepsilon(\varepsilon+4)} \begin{pmatrix} 2+\varepsilon & -2 \\ -2 & 2+\varepsilon \end{pmatrix}$$

$$\overset{F}{X} = \widetilde{A}^{-1} A^{+} \overset{F}{Y} = \frac{1}{\varepsilon(\varepsilon+4)} \begin{pmatrix} 2+\varepsilon & -2 \\ -2 & 2+\varepsilon \end{pmatrix} \begin{pmatrix} y_1 + y_2 \\ y_1 + y_2 \end{pmatrix}$$

$$\overset{F}{X} = \frac{y_1 + y_2}{\varepsilon + 4} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \begin{cases} x_1 = \dfrac{y_1 + y_2}{\varepsilon + 4} \\ x_2 = \dfrac{y_1 + y_2}{\varepsilon + 4} \end{cases}$$

Залежність розв'язків від параметру регуляризації

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$$

$$\varepsilon := 0.1$$

$$\left(A^T \cdot A + \varepsilon \cdot I\right)^{-1} \cdot A^T \cdot y = \begin{pmatrix} 2.439 \\ 2.439 \end{pmatrix}$$

$$\varepsilon := 0.01$$

$$\left(A^T \cdot A + \varepsilon \cdot I\right)^{-1} \cdot A^T \cdot y = \begin{pmatrix} 2.494 \\ 2.494 \end{pmatrix}$$

$$\varepsilon := 0.001$$

$$\left(A^T \cdot A + \varepsilon \cdot I\right)^{-1} \cdot A^T \cdot y = \begin{pmatrix} 2.499 \\ 2.499 \end{pmatrix}$$

$$\left| AX - Y \right| \le \delta$$

$$\varepsilon \le \frac{\left\| A^{+}AA^{+} \right\| \delta}{\left\| A^{+}X(\delta) \right\| - \left\| A^{+} \right\| \delta} \qquad \varepsilon \le \frac{\left\| AA^{+} \right\| \delta}{\left\| X(\delta) \right\| - \delta} \qquad \varepsilon \le \frac{\left\| A \right\|^{2} \delta}{\left\| X(\delta) \right\| - \delta}$$

δ – точність експериментальних даних

Псевдообернені матриці, Псевдорішення

$$A^{-1}A = AA^{-1} = I. \xrightarrow{\text{узагальнення}} \begin{array}{l} A^{\#}AA^{\#} = A^{\#} \\ AA^{\#}A_{,} = A \end{array}$$

$$A^{\#} = \lim_{\varepsilon \to +0} (A^{+}A + \varepsilon I)^{-1} A^{+} \qquad X = A^{\#}Y$$

Оцінки точності рішення матричного рівняння

$$A\vec{X} = \vec{Y}$$

$$\vec{R} = A\vec{X} - \vec{Y}$$ невязка

$$\eta = \vec{R} \cdot \vec{R} = \left\| AX - Y \right\|$$

$$\Delta = X - \widetilde{X}$$ $\widetilde{X}$ Точне рішення

$$A\Delta = A(X - \widetilde{X}) = AX - A\widetilde{X} = AX - Y = R$$

$$A^{-1}A\Delta = A^{-1}(AX - Y) = A^{-1}R$$

$$\Delta = A^{-1}R$$

$$\left\| X - \widetilde{X} \right\| = \left\| A^{-1} \right\| \cdot \| R \|$$

Деякі аспекти обчислення

- Кількість «реальних» чисел обмежена
- Точність зберігання реальних чисел обмежена

Експоненціальна форма запису (нормалізована форма) - представлення дійсних чисел у вигляді мантиси і порядку. Зручна при представлені дуже великих і дуже малих чисел, а також для уніфікації їх написання.

N — число; M — мантиса; n — основа показательной функции; p порядок.

$$N = M \cdot n^p$$

Приклад: $-23.04579 \cdot 10^{-8}$

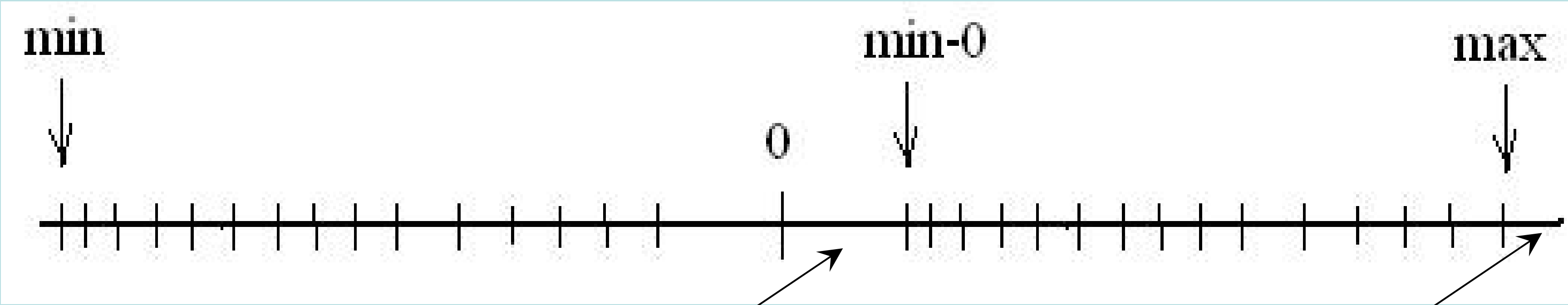
$$-0.\underbrace{2304579}_{\text{мантиса}}\text{E}-06$$

Знак числа

Знак порядку

Порядок числа

«Дійсна шкала» чисел компьютера



Underflow
(зникнення порядку)

Overflow
(переповнення)

Стандарт IEEE754

Одинарная точность
4 байт = 32 біта:

1	8	23
	порядок	мантисса

подвійна точність
8 байт = 64 біта:

1	11	52
	порядок	мантисса

Особлива ситуація - «не число» (NaN).

Точность	Одинарна	Подвійна	Розширена
Розмір (байти)	4	8	10
Число десятинних знаків	7	15	19
Найменше значення > 0	$1{,}2\times10^{-38}$	$2{,}3\times10^{-308}$	$3{,}4\times10^{-4932}$
Найбільше значення	$3{,}4\times10^{+38}$	$1{,}7\times10^{+308}$	$1{,}1\times10^{+4932}$

Не типова (для персональних) комп'ютерів
Надто коштовний режим – витрати часу

Приклади “неточності” обчислень

$$A := \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$$

$$|A| = 0$$

$$B := A^{-1}$$

$$A \cdot B = \blacksquare$$

$$A := \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 8.9 \end{pmatrix}$$

$$|A| = 0.3$$

$$B := A^{-1}$$

$$A \cdot B = \begin{pmatrix} 1.00000000 & \underline{-3.55271368 \times 10^{-15}} & 0.00000000 \\ 0.00000000 & 1.00000000 & 0.00000000 \\ \underline{-1.42247325 \times 10^{-15}} & \underline{3.19189120 \times 10^{-15}} & 1.00000000 \end{pmatrix}$$

$$A := \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 8.99999999 \end{pmatrix} \qquad |A| = 3 \times 10^{-8}$$

$$B := A^{-1}$$

$$A \cdot B = \begin{pmatrix} \underline{0.99999999} & \underline{29.80232239 \times 10^{-9}} & 0.00000000 \\ 0.00000000 & 1.00000000 & 0.00000000 \\ \underline{14.90116119 \times 10^{-9}} & \underline{-29.80232239 \times 10^{-9}} & 1.00000000 \end{pmatrix}$$

Значущі цифри в наближених обчисленнях - всі цифри числа, починаючи з першої зліва, відмінної від нуля, до останньої, за правильність якої можна ручатися.

МАТЕМАТИЧНА ТОЧНІСТЬ РІШЕННЯ має бути вищою (скажімо на порядок !), ніж очікувана фізична точність моделі (експериментальні дані). Вища математична точність, як і нижча (що особливо важливо) неадекватні даної моделі.

Чотири джерела похибки результату розрахунку:

•похибка математичної моделі пов'язана з її невідповідністю фізичної реальності.

•похибка вихідних даних, прийнятих до розрахунку. Це непереборна похибка, але цю похибку можна оцінити для вибору алгоритму розрахунку і точності обчислень.

•похибка методу – заснована на дискретному характері будь-якого чисельного алгоритму. Це означає, що замість точного вирішення вихідної задачі метод знаходить рішення іншого завдання, близького у сенсі до шуканому.

• похибка округлення – пов'язана з використанням у обчислювальних машинах чисел із кінцевою точністю уявлення.

Приклад

$$1 + 10^{20} - 10^{20} = ?$$

$$(1+10^{20}) - 10^{20} = 10^{20} - 10^{20} = 0$$

$$1+10^{20} = 0.000000000000000000001 \cdot 10^{20} + 1 \cdot 10^{20}$$
$$= (0.000000000000000000001+1) \cdot 10^{20}$$

В регістрі с одинарною точність вміщується 7 десятинних цифр !

$$1+ (10^{20} - 10^{20}) = 1 + 0 = 1$$

Приклад

$$e^x = 1 + x + x^2/2 + x^3/3! + \ldots$$

$$e^{-5.5} = 1.0000 - 5.5000 + 15.125 - 27.730 + 39.129 - \ldots..$$

$$= \underline{+0.0026363}$$

Точное значение = <u>+0.00408677</u>

$$e^{-5.5} = 1/e^{5.5} = 1/(1 + 5.5 + 15.125 + 27.730\ldots..)$$

$$= +0.0040865$$

Приклад

$$E_n = \int_0^1 x^n e^{x-1} dx, \quad n = 1,2,3,\ldots$$

$$E_n = 1 - nE_{n-1} \qquad E_1 = 1/e, \qquad n = 1,2,3,\ldots$$

$$E_1 = 0.367879$$

$$E_2 = 0.264242$$

$$\ldots\ldots\ldots$$

$$E_9 = \underline{-0.0684800}$$

$$E_{n-1} = (1 - E_n)/n \qquad E_9 = 0.0916123$$

Математик робить все, що можна і так, як потрібно.
Фізик робить все, що потрібне і так, як можна.
Інженер (хімік) робить все, що потрібне і так, щоб воно працювало.