УДК 519.6: 517

НЕКОРЕКТНІСТЬ ЗАДАЧІ РОЗКЛАДАННЯ ВЕКТОРА ТА НЕРІВНІСТЬ КОШІ – БУНЯКОВСЬКОГО

Проф. Залмен Філер, асист. Олександр Дреєв

Задача разложения вектора по базису подпространства, не вмещающего данный вектор, некорректна. Она заменяется задачей отыскания наилучшего приближения методом наименьших квадратов. Определяется проекция вектора на подпространство и расстояние до него. Попутно получается неравенство Коши – Буняковского.

The problem of decomposition of a vector in basis underspace not containing a vector, is incorrect. She is replaced with a problem the best approximation by a method of the least squares. The projection of a vector on underspace and distance up to him is determined. Inequality Cauchy - Bunyakovsky in passing turns out.

Вступ. Для колінеарних ненульових векторів **a** і **b** рівняння **a** = λ**b** має єдиний розв'язок відносно λ. У евклідовому просторі його можна знайти, множачи скалярно рівняння на вектор **b**. Це дає $\lambda=(\mathbf{a},\mathbf{b})/b^2$. Для не колінеарних векторів указане рівняння не має розв'язків. Задача розв'язання цього рівняння є *некоректною* (тобто, розв'язок або не існує, або не єдиний, або "дуже чутливий" до малих змін умов) й через те, що *мала* зміна напрямку одного з колінеарних векторів приводить її до задачі, яка не має розв'язку. Тому її треба переформулювати так, щоби вона мала єдиний розв'язок завжди. Такого типу задачі виникають й в математичній статистиці, коли навіть для лінійного детермінованого зв'язку між масивами – векторами $\mathbf{a}(a_1, a_2,.., a_n)$ та $\mathbf{b}(b_1, b_2,\ldots, b_n)$ рівняння **a** = λ**b** не має розв'язків, бо теоретично пропорційні елементи $a_к$ та $b_к$ через наявність похибок вимірювання чи обчислення стають непропорційними. Для таких задач К.Гауссом та А.Лежандром був запропонований метод найменших квадратів, де шукається значення λ, яке мінімізує квадратичну нев'язку - відхил $(\mathbf{a} - \lambda\mathbf{b}, \mathbf{a} - \lambda\mathbf{b})$.

1. Знаходження мінімальної нев'язки. Ми вже писали про доведення нерівності Коші – Буняковського за допомогою переходу до ортів векторів і дослідження нерівностей для модуля суми або різниці ортів. В цій статті ми використаємо дослідження функції – скалярного добутку

$$f(\lambda)=(\mathbf{a} - \lambda\mathbf{b}, \mathbf{a} - \lambda\mathbf{b}) = (\mathbf{a} - \lambda\mathbf{b})^2. \qquad (1)$$

За властивостями скалярного добутку $f(\lambda)>0$ при всіх **a, b** і λ й дорівнює нулеві тільки при $\mathbf{a} - \lambda\mathbf{b} = \mathbf{0}$, коли $\mathbf{a} = \lambda\mathbf{b}$. Таким чином, мінімум $f(\lambda) = 0$ при таких **a, b** і λ. Зараз ми пропонуємо дослідити функцію $f(\lambda)$ на мінімум й при не колінеарних векторах **a, b**.

Розглянемо найпростіший випадок евклідового простору для векторів **a, b** й скаляра $\lambda\in R$. Тоді, використовуючи та дистрибутивність скалярного добутку відносно різниці векторів та однорідність відносного скалярного множника, отримаємо

$f(\lambda) = \lambda^2 b^2 - 2\lambda(\mathbf{a},\mathbf{b}) + a^2 \Rightarrow f'(\lambda) = 2\lambda b^2 - 2(\mathbf{a},\mathbf{b})=0$, якщо $\lambda=\lambda_0=(\mathbf{a},\mathbf{b})/ b^2$. Тоді мінімальне значення функції $f(\lambda) = f(\lambda_0) = a^2 - (\mathbf{a},\mathbf{b})/ b^2 = (a^2 b^2 - (\mathbf{a},\mathbf{b})^2) / b^2\geq 0$.

Звідси й випливає нерівність Коші Буняковського a^2b^2 - $(\mathbf{a},\mathbf{b})^2 \geq 0$. Для векторів на площині або у тривимірному просторі вектор **a** - λ_0**b** перпендикуляр до вектора **b**.

В ермітнитовому (унітарному) просторі з комплексними значеннями скалярного добутку **(a,b)** й числа λ, але невід'ємними значеннями скалярного квадрату – функції $f(\lambda)$ у (1), використовуючи антикомутативність скалярного добутку $(\mathbf{a},\mathbf{b}) = (\mathbf{b}, \mathbf{a})^*$ (* - знак комплексного спряження) та винесення спряженого множника другого вектора у добутку $(\mathbf{a}, \lambda\mathbf{b}) = \lambda^*(\mathbf{a},\mathbf{b})$, отримаємо рівність $f(\lambda) = |\lambda|^2b^2 - \lambda^*(\mathbf{a},\mathbf{b}) - \lambda(\mathbf{a},\mathbf{b})^* + a^2$. Підставивши сюди комплексне значення $\lambda=\alpha+i\beta$, отримаємо $f(\alpha,\beta) = (\alpha^2+\beta^2)b^2 - (\alpha-i\beta)(\mathbf{a},\mathbf{b}) - (\alpha+i\beta)(\mathbf{a},\mathbf{b})^* + a^2$. Частинні похідні f_α та f_β у точці мінімуму повинні дорівнювати нулеві, що дає критичні значення α_0 та β_0: $\alpha_0 = \mathrm{Re}((\mathbf{a},\mathbf{b}))/b^2$, $\beta_0 = -\mathrm{Im}(\mathbf{a},\mathbf{b}))/b^2 \Rightarrow \lambda_0=(\mathbf{a},\mathbf{b})^*/b^2$. Тоді мінімальне значення функції $f(\lambda)=f(\lambda_0) = a^2 - |(\mathbf{a},\mathbf{b})|^2/ b^2 =(a^2b^2-|(\mathbf{a},\mathbf{b})|^2)/b^2$. Це дає ту ж саму нерівність Коші – Буняковського $|(\mathbf{a},\mathbf{b})|^2 \leq a^2b^2$. Рівність тут буде для комплексних векторів, коли комплексний множник λ зробить ці вектори "колінеарними". Пояснимо прикладом: нехай маємо вектори **a (**3,2**)** та **b(**4,-1). Вони не є колінеарними, бо їх координати не пропорційні, але, розглядаючи їх як комплексні *числа* **a =** 3+2i та **b =** 4-1i, знайдемо λ = (3+2i)/(4-i) = (3+2i)/(4+i)/17 = (10+11i)/17. Таким чином, λ = 10/17+i11/17 – єдиний розв'язок рівняння в **C**. Він дає коефіцієнт розтягування $|\lambda|$ вектора **b** та кут його повороту $\arg(\lambda)$ для отримання вектора **a**.

2. Елементарний підхід без використання диференціального числення. У підручниках алгебри зазвичай розглядається той самий квадратний тричлен $f(\lambda) = \lambda^2b^2 - 2\lambda(\mathbf{a},\mathbf{b}) + a^2$ й використовується шкільне знання, що він приймає невід'ємні значення, а тому його дискримінант $4(\mathbf{a},\mathbf{b})^2-4a^2b^2 \leq 0$. Такий самий підхід використано в брошурі [1, с. 4]. Там для доведення пропонується розглянути функцію $y = \sum_{i=1}^{n}(xa_i + b_i)^2$. Чому? Такий само штучний підхід уміщено в усіх підручниках лінійної алгебри та функціонального аналізу. Між тим, наш підхід пов'язаний із загальною ідеєю побудови регуляризуючого оператора для розв'язання некоректних задач, а мінімальна нев'язка вказує на відхилення отриманого "псевдо розв'язку" від неіснуючого розв'язку. Точніше, ми отримаємо абсолютну квадратичну похибку $f(\lambda_0)=\dfrac{a^2b^2 - |(a,b)|^2}{b^2}$; якщо вектори **a** і **b** вважати рівноправними, то доцільно оцінювати точність відношенням

$$\delta(\mathbf{a}, \mathbf{b}) = \frac{a^2b^2 - |(a,b)|^2}{a^2b^2}. \quad (2)$$

Тут нев'язку дано й для комплексно-значних векторів в ермітовому просторі, де скалярний добуток є комплекснозначним; його модуль, як і модулі векторів **a** і **b,** є *дійсними* числами.

У посібнику [2, с. 90, 105] пропонується доведення нерівності К.-Б. за допомогою підрахунку f(λ) при $\lambda=\lambda_0=(\mathbf{a},\mathbf{b})^*/b^2$. *Чому* береться таке значення, автор не вказує. Про справедливість нерівності К.-Б. в *унітарному* просторі сказано, що "доведення проводиться за тією ж схемою, як і в дійсному випадку". Нагадаємо, що перше видання цієї книги вийшло у 1974 р., коли в масовій школі ще не вивчали диференціального числення. Та й мети дати мінімізуюче тлумачення указаного значення $\lambda_0=(\mathbf{a},\mathbf{b})^*/b^2$, автор не ставить.

Наш підхід переноситься на векторні рівняння з багатьма (n) скалярними невідомими $x_к$ типу

$$\mathbf{b}_1x_1+ \mathbf{b}_2x_2+ \mathbf{b}_3x_3+\ldots+ \mathbf{b}_nx_n=\mathbf{a}. \qquad (3)$$

Якщо вектор **a** не лежить в оболонці, натягнутій на вектори $\mathbf{b}_1$, $\mathbf{b}_2$,..., $\mathbf{b}_n$, то задача не має розв'язку $\mathbf{x}(x_1, x_2, \ldots, x_n)$. Регуляризація задачі досягається заміною розв'язання несумісної системи (3) задачею знаходження вектора **x**, який *мінімізує* сумарне квадратичне відхилення (нев'язку) $f(\mathbf{x}) = (B\mathbf{x} - \mathbf{a})^2$. Це дає систему *нормальних* рівнянь, яка для лінійно незалежних рівнянь має єдиний розв'язок. При n=2 й не колінеарних $\mathbf{b}_1$, $\mathbf{b}_2$ нев'язкою є квадрат відстані кінця вектора **a** від площини векторів $\mathbf{b}_1$, $\mathbf{b}_2$, тобто квадрат довжини перпендикуляра **c**, опущеного на вказану площину, бо умови – рівняння для пошуку x_1 та x_2, отримані диференціюванням по x_1 та x_2,

$$f_{x1}(x_1, x_2) = 2(x_1\mathbf{b}_1+ x_2\mathbf{b}_2 - \mathbf{a})\mathbf{b}_1=0,\ f_{x2}(x_1, x_2) = 2(x_1\mathbf{b}_1+ x_2\mathbf{b}_2 - \mathbf{a})\mathbf{b}_2=0,$$

є умовами перпендикулярності вектора - нев'язки до векторів $\mathbf{b}_1$ та $\mathbf{b}_2$. При n>2 інші умови теж є умовами перпендикулярності вектора - нев'язки до векторів "базису" $\mathbf{b}_1, \mathbf{b}_2,\ldots, \mathbf{b}_n$: $(x_1\mathbf{b}_1+ x_2\mathbf{b}_2 + \ldots+ x_n\mathbf{b}_n- \mathbf{a})\mathbf{b}_к=\mathbf{0}$, k= 1, 2…,n. В [1, с. 74-76] такий підхід проведено для кожної координати $x_к$, даючи їй приріст t, отримуючи квадратичну функцію від цього t, що веде до нормальної системи. Методами диференціального числення функції багатьох змінних, ці рівняння отримуємо миттєво.

Якщо вектори $\mathbf{b}_1, \mathbf{b}_2,\ldots, \mathbf{b}_n$ взаємно ортогональні, то для кожного з $x_к$ маємо одне рівняння $(\mathbf{b}_к, \mathbf{b}_к)x_к = (\mathbf{a}, \mathbf{b}_к)$, звідки $x_к = (\mathbf{a}, \mathbf{b}_к)/ b_к^2$.

У [1, с. 81] наводиться формула Гаусса для мінімального квадратичного відхилення $f(\mathbf{x}) = (B\mathbf{x} - \mathbf{a})^2$ при (без виведення) при n = 3, але не вказується можливість *узагальнення* з неї нерівності К.-Б.

Величина нев'язки більш інформативна, ніж нерівність Коші – Буняковського. Формула (2) показує на скільки він відрізняється відносно добутку модулів цих векторів. Він оцінює коефіцієнт К кореляції між векторами – масивами **a** та **b:**

$$\delta(\mathbf{a},\mathbf{b}) = \frac{a^2b^2 - |(a,b)|^2}{a^2b^2} = 1 - К,\ К = \frac{(a,b)^2}{a^2b^2} \text{ в } \mathbf{R}. \qquad (3)$$

Якщо вектори **a** і **b** ***колінеарні***, тобто відповідні координати пропорційні, то К=1. При неколінеарності цих векторів за нерівністю Коші – Буняковського чисельник у виразі для δ(**a,b) додатний** й коефіцієнт К<1. Для незалежних випадкових значень елементів масивів **a** і **b** скалярний добуток (**a, b) ≈ 0 й К≈0**, що свідчить про відсутність суттєвого зв'язку між **a** та **b**.

Сучасні програмні засоби дозволяють автоматизувати розрахунки з використанням векторів та матриць. Наприклад, використовуючи MatLab, були згенеровані випадкові матриця **B** розмірністю 4^x2 та вектор стовпець **a** розмірністю 4 (B=rand(4,2); **a**=rand(4,1)):

$$\mathbf{a}^{T}=(0.8214;\ 0.4447;\ 0.6154;\ 0.7919), B^{T}=\begin{pmatrix}0.9501 & 0.2311 & 0.6068 & 0.4860\\ 0.8913 & 0.7621 & 0.4565 & 0.0185\end{pmatrix}.$$

Отримано за допомогою виразу $\mathbf{x}=(B^TB)^{-1}B^T\mathbf{a}=(1.0697;\ -0.0212)^T$. Обчислено величину К= (B*x-a)'*(B*x-a)/(a'*a*((B*x)'*(B*x)))= 0.0467. Мале значення К можна (порівняно з одиницею) свідчить про відсутність тісного зв'язку вектора **a** зі стовпцями матриці **B**.

При дослідженнях залежних величин коефіцієнт К матиме значення близьке до 1. Як приклад було розраховано коефіцієнт кореляції К для вектора **a** – чисел Вольфа та матриці В, стовпці якої із запізненням на 8 діб містять швидкість сонячного вітру, його густину та значення напруженості міжпланетного магнітного поля (за 100 діб). Коефіцієнт К=0.7913, що свідчить про суттєвий зв'язок між цими величинами.

3. Лінійні регресії a на b та b на a. Якщо вектори **a** та **b** не колінеарні, то їх можна представити у вигляді $\mathbf{a} = \lambda\mathbf{b} + \mathbf{a}_0$ або $\mathbf{b} = \mu\mathbf{a} + \mathbf{b}_0$. Логічно, що пошук λ та μ з умов мінімізації модуля вектора **a - λb** чи **b - μa,** дасть ті ж значення $\lambda_0 = (\mathbf{a},\mathbf{b})/\ b^2$, $\mu_0 = (\mathbf{a},\mathbf{b})/a^2$, які дає мінімізація модулів векторів - нев'язок. Очевидно, $\lambda_0\mu_0 = (\mathbf{a},\mathbf{b})^2/(a^2b^2)=K\in[0;1]$. Ці числа називають відповідно *коефіцієнтами регресії* **a** на **b** та **b** на **a**. Вони мають однаковий знак "+" при (**a,b)**>0 та **"-"** при (**a,b)**<0. Для геометричних векторів **a** і **b** гострий кут між ними дасть знак "+", тупий – знак **"-".** Вектори нев'язок $\mathbf{a}_0$ та $\mathbf{b}_0$, як правило, різні й ортогональні відповідно векторам **b** та **a**.

Висновки

1. Аналізуються доведення нерівності Коші – Буняковського (К.-Б.) засобами алгебри та аналізу. Відмічений штучний характер цих методів.
2. Пропонується йти від квадратичного наближення одного вектора колінеарним другому з умови мінімізації нев'язки (відхилу) в евклідовому векторному просторі. Отримані відповідні коефіцієнти, які є коефіцієнтами регресії одного вектора – масиву на другий.
3. Метод і результати переносяться на вектори в ермітовому просторі.
4. Показано, що результати виражають загальну лінійну залежність між векторами виду $a = \lambda b + a_0$ чи $b = \mu a + b_0$.

5. Дано статистичне тлумачення результатів. Показано за допомогою нерівності К.–Б., що коефіцієнт кореляції $(a,b)^2/(a^2\ b^2)=:K\in[0;1]$.
6. Підходи, доступні студентам 1-го курсу ВНЗ, які вивчають вищу математику.
7. Існування розвинутого програмного забезпечення дозволяє, не проводячи складних аналітичних перетворень, отримувати вектор x, що мінімізує квадратичне відхилення $(a\text{-}Bx)^2$, та відносну оцінку мінімального відхилення наближення від заданого вектору a.

БІБЛІОГРАФІЯ

1. Дринфельд Г.И. Интерполирование и способ наименьших квадратов. – К.: Вища школа, 1984. – 103 с.
2. Воеводин В.В. Линейная алгебра. - М.: Наука, 1980. – 400 с.
3. Дрінфельд Г.І. Вища математична освіта у Києві в роки 1927-1941// У світі математики, т.4, 1998,вип. 2. – С. 68 – 82.
4. Карнацевич В.Л. 100 знаменитых харьковчан. – Харьков: Фолио, 2005. – 510 с.
5. Алексеев Е.Р., Чеснокова О.В. Решение задач вычислительной математики в пакетах MahCad 12, MATLAB 7, Maple 9. – М.: NT Press, 2006. – 496 с. (Самоучитель)