## **Багатовимірний регресійний аналіз.**

### **Кореляції вищих порядків**

### 

У попередніх викладках ми розглядали кореляцію між компонентами двовимірного вектора. Очевидно, такі ж питання виникають, коли розглядаємо випадкові вектори з кількістю координат більшою двох.

Нехай випадковий вектор має координати (). Хочемо вивчити кореляцію, наприклад, між  та . Для цього будемо вважати координати  зафіксованими, тобто такими, які набули певних значень з відповідних їх областей визначення і далі не змінюються. Однак, проблема полягає в тому, що кореляція між виділеними для аналізу змінними  та  може бути обумовлена тим, що обидві вони цілком, чи частково залежать, наприклад, від третьої змінної , в той час як при кожному фіксованому значенні цієї змінної  змінні  та  стохастично незалежні. Тому спочатку розглянемо кореляцію між  та  тривимірного випадкового вектора ().

Вплив третьої координати на перші дві спробуємо нейтралізувати замінивши змінні  та  такими величинами , , коефіцієнти кореляції яких з  рівні нулю. Якщо , то коефіцієнт кореляції

. (1)

Отже множники  потрібно вибирати так, щоб

, .

Звідси

, (2)

. (3)

Коефіцієнт кореляції між  та  без впливу змінної  дорівнює

. (4)

Розкривши дужки та використовуючи відповідні властивості математичного сподівання, формулу (1), переконуємося, що

 =

=

-.

Підставивши в цю рівність формули (2), (3), одержимо:

=

=(. (5)

Аналогічно доводимо, що

, (6)

. (7)

Підставивши (5), (6), (7) в (4), остаточно маємо:

. (8)

Використовуючи замість випадкових змінних  їх вибіркові значення відповідно , цілком так само, як отримано формулу (8), отримуємо співвідношення

, (9)

де  вибіркові кореляції, відповідно, між першою і другою координатою, між першою та третьою, між другою та третьою координатою випадкового вектора (),  умовна кореляція між першою та другою координатою цього вектора при зафіксованій третій координаті.

Формулу (9) узагальнюємо індукцією за розмірністю простору випадкових змінних наступним чином. Індекси, які відповідають постійним значенням змінних, будемо називати ***німими***, а кількість німих індексів – ***порядком частинної кореляції****.* Німі індекси записуватимемо після відкриваючої квадратної дужки. Тоді, наприклад,  ***частинна вибіркова кореляція*** () – го порядкуміж координатами  та . Визначимо рекурентно частинну вибіркову кореляцію () – го порядку через три частинні вибіркові кореляції () – го порядку за формулою:

. (10)

**Приклад**. Частинна кореляція між першою і другою координатою тривимірного вектора, згідно (10), як і згідно формули (9), рівна:

,

а кореляція між тими ж координатами чотиривимірного вектора:–

 . ■

З означення кореляцій вищого порядку випливає , що ***активн****і,*  незафіксовані індекси можна переставляти місцями, а зафіксовані, німі, ***пасивні*** переставляти не можна – відносно них кореляція несиметрична.

Отже маємо цілий набір частинних кореляцій, які визначаються різними перестановками індексів в них. Щоб обчислити частинні кореляції  порядку необхідно мати всі такі частинні кореляції нульового порядку:



### 

### **Варіанси вищих порядків**

### 

***Варіанси вищих порядків*** також будуємо за аналогією до формування кореляцій вищих порядків, тобто фіксуванням всіх неактивних координат випадкового вектора. Отже така варіанса має один активний індекс, а інші () – пасивні, німі. Цю ***вибіркову частинну варіансу*** будемо називати ***варіансою* (****) *– го порядку***. Варіансу () –го порядку визначаємо рекурентно через варіансу () – го порядку та кореляцію () – го порядку за формулою:

. (11)

**Приклад.** Частинна варіанса першої координати двовимірного вектори при фіксованій другій координаті, згідно (11), дорівнює:

,

а частинна варіанса першої координати тривимірного вектори при фіксованих двох інших:

. ■

Послідовно застосовуючи формулу (11), можна виразити варіансу  – го порядку через варіансу першого порядку:

 . (12)

***Стандарт m-ого порядку*** визначають звичайним чином:

 . (13)

### **Регресії вищих порядків**

Регресії вищих порядків також визначаємо за індукцією. Позначимо вибіркову регресію між  та  при постійних значеннях всіх інших координат  через . Тоді ***частинну регресію  - порядку*** визначають через кореляцію ( ) – го порядку та два стандарти  – го порядку за формулою:

. (14)

Зокрема, при  одержуємо регресію першого порядку

,

а при =2 – відому вже раніше регресію нульового порядку

.

### 

### **Лінійне рівняння регресії однієї змінної на інші**

За аналогією до двовимірного випадку, можна говорити і про лінійне рівняння регресії однієї координати вимірного випадкового вектора () відносно всіх інших координати цього ж вектора. Очевидно, що вивід таких рівнянь буде подібним до двовимірному випадку з урахуванням того, що тепер потрібно аналізувати частинні регресії вищих порядків.

Нехай в результаті спостережень над випадковим вектором () одержано  різних значень (), () його, причому значення  цей вектор набував  разів. Очевидно, що .Обчислимо середні арифметичні



координат вектора (). Будемо шукати, наприклад, гіперплощину

, (15)

яка проходить через точку () і найкращим чином, в сенсі принципу найменших квадратів (див. § 5), узгоджується з експериментальними даними , (). Отже задача полягає у відшуканні таких коефіцієнтів , при яких сума



набуває найменшого значення. З необхідної умови мінімуму цієї суми маємо систему рівнянь:

== 0, (), (16)

розв’язавши яку з (15) отримаємо лінійне рівняння регресії першої компоненти - мірного випадкового вектора на всі інші координати:



. (17)

При  це рівняння перетворюється у відоме рівняння прямої регресії першої координати двовимірного випадкового вектора на другу координату.

Емпіричне рівняння регресії *m* - го порядку аналітично зображає випадковий процес, що описується випадковими змінними  і слугує для подання великого за обсягом статистичного матеріалу компактно аналітично.