$$\mathcal{E}_2$$

Э П С И Л О Н

E P S I L O N

журнал об эконометрике
и не только о ней

Корреляция: простая, частная и условная

Борис Демешев*

5 августа 2015 г.

Аннотация

Корреляция — это способ описать силу линейной зависимости между двумя случайными величинами одним числом. Каков геометрический смысл корреляции? Что такое частная корреляция? Как связаны частная и условная корреляция?

Ключевые слова: корреляция, частная корреляция, условная корреляция, косинус, проекция.

1 Сколько вешать в граммах?

Почему мы измеряем температуру тела с помощью градусника?

1. Измерить температуру очень удобно
2. Это измерение несёт в себе информацию о здоровье

Сложное описание здоровья сводится измерением температуры к одной цифре. Естественно, куча информации теряется в этой цифре и бессмысленно лечить человека, руководствуясь только температурой его тела. Температура 39° говорит, что не так, но что — непонятно. А температура 36.6° ещё не говорит о том, что у человека идеальное здоровье. Однако процедура очень проста и в некоторых ситуациях (например, при обыкновенной простуде) её достаточно для принятия решения о приёме жаропонижающего.

Если бы для измерения температуры нужно было специальное устройство размером с половину комнаты, никто бы дома её не мерял. Простота измерения очень важна!

Подобном образом дела обстоят и с описанием зависимости между случайными величинами. Зависимость между случайными величинами полностью описывается их совместной функцией распределения, $F(x, y) = ¶(X \leq x, Y \leq y)$. Вместо сложной функции распределения мы хотим получить одно число. Некую «силу зависимости». Назовём это мифическое число $Dep(X, Y)$.

Что мы требуем от этого числа?

1. Посчитать это число очень удобно
2. Это число несёт в себе информацию о зависимости

*НИУ ВШЭ, Москва.

В каком смысле «удобно» считать?

Очень часто возникают суммы случайных величин, поэтому было бы здорово, чтобы у суммы легко считалась наша характеристика $Dep(X, Y)$. Проще всего было бы, если бы:

$$Dep(X + Z, Y) = Dep(X, Y) + Dep(Z, Y)$$

И, конечно, мы ждём, что у независимых случайных величин нулевая сила зависимости, $Dep(X, Y) = 0$, а ненулевая сила зависимости, $Dep(X, Y) \neq 0$, была бы возможна только у зависимых случайных величин.

Этим двум требуемым свойствам (простота подсчёта и информация о зависимости) отвечает ковариация.

Определение 1. *Ковариация величин X и Y измеряет ...*

$$\mathrm{Cov}(X, Y) = \mathbb{E}\left[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right] = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

здесь рассказать про прямоугольники и площадь с плюсом/минусом?

2 Корреляция по-русски

Обычно в учебниках даётся такое определение корреляции

$$\mathrm{Corr}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}. \tag{2.1}$$

Естественно, возникает вопрос: «С какого перепугу? Почему это мы делим ковариацию на что-то там?»

Мы дадим определение корреляции словами:

Определение 2. *Корреляция между случайными величинами X и Y показывает на сколько своих стандартных отклонений в среднем растёт случайная величина Y при росте случайной величины X на одно своё стандартное отклонение.*

А теперь из этого словесного определения мы получим формулу 1.1. Разложим величину Y на два слагаемых. Первое слагаемое вбирает в себя всю ту часть Y, которая линейно зависит от X, а второе — всё оставшееся:

$$\frac{Y}{\sigma_Y} = \rho \cdot \frac{X}{\sigma_X} + \varepsilon$$

В этой формуле видно, что с ростом X на одно стандартное отклонений σ_X правая часть изменится в среднем на ρ, и, следовательно, величина Y в среднем изменится на $\rho \cdot \sigma_Y$.

Естественно, мы хотим, чтобы с ростом X величина ε в среднем не менялась, то есть хотим нулевую «силу зависимости» между ними, $\mathrm{Cov}(X, \varepsilon) = 0$.

$$\mathrm{Cov}\left(X, \frac{Y}{\sigma_Y} - \rho \cdot \frac{X}{\sigma_X}\right) = 0$$

По свойствам ковариации получаем

$$\mathrm{Cov}(X,Y)/\sigma_Y - \rho\,\mathrm{Cov}(X,X)/\sigma_X = 0$$

И, тадам, выражаем корреляцию, ρ:

$$\rho = \frac{\mathrm{Cov}(X,Y)/\sigma_Y}{\mathrm{Cov}(X,X)/\sigma_X} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y}$$

Несмотря на асимметричность исходного разложения (эпсилон прибавляется в правой части уравнения к величине X), результирующая формула для корреляции получается симметричной. Из этого следует, что ровно такой же результат получится, если начать с разложения:

$$\frac{X}{\sigma_X} = \rho \cdot \frac{Y}{\sigma_Y} + \varepsilon$$

Из определения неочевидно, что корреляция лежит в пределах от -1 до 1. ...

Стоит обратить внимание на немного контр-интуитивный факт. Если бы зависимость между X и Y была бы жесткой детерминистической, и с ростом X на единицу величина Y росла бы на Δ, то с ростом Y на единицу величина X росла бы на $1/\delta$. Для случайных величин обращения не происходит. Если с ростом X на одно своё стандартное отклонение величина Y в среднем растёт на ρ своих стандартных отклонений, то и с ростом Y на одно своё стандартное отклонение величина X в среднем растёт на ρ своих стандартных отклонений.

? парадокс возвращения к среднему ?

3 Геометрический смысл корреляции

Давайте рисовать случаные величины векторами-стрелочками! Не в том смысле, что у стрелочки случайное направление или длина, а в том смысле, что направление и длина стрелочки описывают характеристики этой случайной величины.

Любую геометрию можно задать, задав скалярное произведение. Действительно, если мы умеем считать скалярное произведение двух любых векторов, $<\vec{a},\vec{b}>$, то длина вектора считается ровно как в 9-м классе:

$$|\vec{a}| = \sqrt{<\vec{a},\vec{a}>}$$

И также любой девятикласник помнит, что косинус угла между векторами считается как

$$\cos(\vec{a},\vec{b}) = \frac{<\vec{a},\vec{b}>}{|\vec{a}||\vec{b}|}$$

Мы определим скалярное произведение двух случайных величин как их ковариацию:

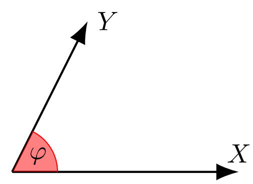
$$<X,Y>= \mathrm{Cov}(X,Y)$$

При таком подходе длиной случайной величины окажется стандартное отклонение:

$$\sqrt{\mathrm{Cov}(X,X)} = \sqrt{\mathrm{Var}(X)} = \sigma_X$$

А корреляция окажется косинусом угла между случайными величинами:

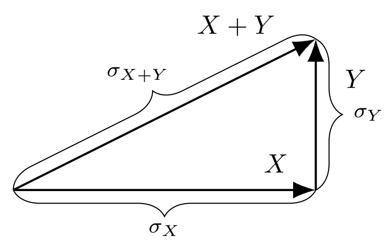
$$\cos\varphi = \cos(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y} = \mathrm{Corr}(X, Y)$$



Значит в нашей геометрии длина стрелочки — стандартное отклонение случайной величины, а косинус угла между двумя стрелочками — это корреляция двух случайных величин. Дисперсия, следовательно, это квадрат длины случайной величины. Перпендикулярными случайными величинами будут те, косинус угла между которыми равен нулю, то есть некоррелированные.

Например, сформулируем в данной геометрии теорему Пифагора. Если случайные величины X и Y перпендикулярны (корреляция или ковариация равны нулю), то дисперсия их суммы (квадрат длины гипотенузы) равен сумме их дисперсий (сумму квадратов длин катетов):

$$\mathrm{Var}(X + Y) = \sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y = \mathrm{Var}(X) + \mathrm{Var}(Y)$$



Введение геометрии позволяет говорить о проекции. Например, можно спроецировать случайную величину Y на множество случайных величин пропорциональных величине X, $\{cX | c \in \mathbb{R}\}$. Если на обычной плоскости спроецировать вектор $\vec{a}$ на прямую, порожденную вектором $\vec{b}$, то получится $\cos(\vec{a}, \vec{b}) \cdot \vec{b}$. По аналогии, если спроецировать случайную величину Y на множество $\{cX | c \in \mathbb{R}\}$, то получится $\mathrm{Corr}(X, Y) \cdot X$. Другими словами, среди случайных величин пропорциональных X величина $\hat{Y} = \mathrm{Corr}(X, Y) \cdot X$ — самая похожая на величину Y.

Понятие проекции позволяет интепретировать квадрат корреляции. Квадрат косинуса равен отношению квадрата длины прилежащего катета $\mathrm{Var}(\hat{Y})$ к квадрату гипотенузы $\mathrm{Var}(Y)$.

(картинка)

Следовательно, $\mathrm{Corr}(X, Y)^2 = \frac{\mathrm{Var}(\hat{Y})}{\mathrm{Var}(Y)}$, то есть квадрат корреляции показывает долю дисперсии Y, которую можно объяснить с помощью величин пропорциональных X.

4 Корреляция и независимость

Теорема 1. *Случайные величины X и Y независимы тогда и только тогда, когда некоррелированы любые функции $f(X)$ и $g(Y)$.*

Другими словами для независимости X и Y необходима некоррелированность пар X и Y, X^2 и $\cos(Y)$, $\exp(X)$ и $1/Y$, и так далее. Из этого следует, что некоррелированность X и Y является необходимым, но недостаточным условием для независимости.

Можно выделить три «степени» независимости случайных величин X и Y:

Некоррелированность Y и X	$\operatorname{Corr}(X, Y) = 0$
$\mathbb{E}(Y\|X) = \mathbb{E}(Y)$	$\operatorname{Corr}(f(X), Y) = 0$ для всех $f()$
Независимость Y и X	$\operatorname{Corr}(f(X), g(Y)) = 0$ для всех $f()$ и $g()$

Многие ошибочно считают, что если величина X имеет нормальное распределение $N(\mu_X, \sigma_X^2)$ и величина Y имеет нормальное распределение $N(\mu_Y, \sigma_Y^2)$, и X и Y некоррелированы, то они независимы. Это неверно.

Контрпример. Случайная величина X имеет стандартное нормальное распределение $N(0; 1)$, случайная величина Z независима от X и равновероятно принимает значения -1 и $+1$. Определим величину Y как их произведение, $Y = XZ$.

В этом примере величины X и Y зависимы, так как $|X| = |Y|$. Однако Y распределена нормально стандартно и $\operatorname{Corr}(X, Y) = 0$.

Правильная теорема звучит так:

Теорема 2. *Если некоррелированные случайные величины X и Y имеют совместное нормальное распределение, то X и Y независимы.*

Попутно упомянем ещё одно неожиданное свойство предъявленного контрпримера. Если случайные величины нормальны по отдельности, то вполне возможно, что их сумма ненормальна. Для пары величин, имеющих совместное нормальное распределение, это невозможно.

5 Частная корреляция

Определение 3. *Частная корреляция между величинами X и Y при фиксированной величине Z показывает на сколько своих стандартных отклонений σ_Y в среднем вырастет Y при росте величины X на одно своё стандартное отклонение σ_X и постоянном значении величины Z.*

Для нахождения частной корреляции используется разложение

$$\frac{Y}{\sigma_Y} = \rho_{XY|Z} \cdot \frac{X}{\sigma_X} + \rho_{YZ|X} \frac{Z}{\sigma_Z} + \varepsilon$$

Альтернативный подход к подсчёту частной корреляции следующий:

1. Спроецируем X на множество величин, некоррелированных с Z. Получим $\tilde{X}$.
2. Спроецируем Y на множество величин, некоррелированных с Z. Получим $\tilde{Y}$.
3. Частная корреляция между X и Y при фиксированной Z — это обычная корреляция между $\tilde{X}$ и $\tilde{Y}$.

(картинка ...)

Два подхода к определению частной корреляции эквивалентны в силу теоремы Фриша-Ву-Ловелла (Frisch–Waugh–Lovell). Обычно эта теорема формулируется применительно к регрессии, а здесь мы приведём её вариат для случайных величин.

Теорема 3. *Если имеют место разложения:*

$$Y = a_1 Z_1 + a_2 Z_2 + \ldots + a_n Z_n + \tilde{Y}, \text{ где } \tilde{Y} \perp Z_1, Z_2, \ldots, Z_n$$

и

$$X = b_1 Z_1 + b_2 Z_2 + \ldots + b_n Z_n + \tilde{X}, \text{ где } \tilde{X} \perp Z_1, Z_2, \ldots, Z_n$$

То в разложениях

$$\tilde{Y} = d\tilde{X} + \varepsilon, \text{ где } \varepsilon \perp \tilde{X}$$

и

$$Y = c_1 Z_1 + c_2 Z_2 + \ldots + c_n Z_n + dX + u, \text{ где } u \perp Z_1, Z_2, \ldots, Z_n, X$$

коэффициенты при $\tilde{X}$ и X совпадают.

6 Условная корреляция

Определение 4. *Условная корреляция между величинами X и Y при известном значении величины Z показывает на сколько своих стандартных отклонений σ_Y в среднем вырастет Y при росте величины X на одно своё стандартное отклонение σ_X при заданном значении величины Z.*

Следует подчеркнуть одно существенное отличие условной корреляции от обычной и частной. Обычная и частная корреляция являеются константами. Условная корреляция $\operatorname{Corr}(X, Y|Z)$ является функцией от Z. Величина Z является случайной, поэтому и условная корреляция $\operatorname{Corr}(X, Y|Z)$ является случайной величиной.

Здесь регрессионное определение????

Чуть более формальное определение:

Определение 5.

$$\operatorname{Corr}(X, Y|Z) = \frac{\operatorname{Cov}(X, Y|Z)}{\sqrt{\operatorname{Var}(X|Z)\operatorname{Var}(Y|Z)}},$$

где $\operatorname{Cov}(X, Y|Z) = \mathbb{E}(XY|Z) - \mathbb{E}(X|Z)\mathbb{E}(Y|Z)$ и $\operatorname{Var}(X|Z) = \mathbb{E}(X^2|Z) - (\mathbb{E}(X|Z))^2$

Пример подсчета частной и условной корреляций.

Пример 1.

Закон распределения случайных величин X_1, X_2, X_3 задан двумя таблицами:

$X_3 = 0$

	$X_2 = 0$	$X_2 = 1$
$X_1 = 0$	0.06	0.08
$X_1 = 1$	0.24	0.32

$X_3 = 1$

	$X_2 = 0$	$X_2 = 1$
$X_1 = 0$	0.1	0
$X_1 = 1$	0	0.2

Найдите условную корреляцию $\mathrm{Corr}(X_1, X_2|X_3)$ и частную корреляцию $\mathrm{pCorr}_{X_3}(X_1, X_2)$.

Решение.

Эти две таблички на самом деле реализуют простую мысль: при $X_3 = 1$ величины X_1 и X_2 связаны детерминистически линейно, а при $X_3 = 0$ величины X_1 и X_2 независимы.

Считаем две вспомогательные условные корреляции, $\mathrm{Corr}(X_1, X_2|X_3 = 0) = 0$, $\mathrm{Corr}(X_1, X_2|X_3 = 1) = 1$.

Отсюда получаем, что $\mathrm{Corr}(X_1, X_2|X_3) = X_3$. Для дискретных случайных величин запись условного ожидания не однозначна, и, например, ответ $\mathrm{Corr}(X_1, X_2|X_3) = X_3^2$ также будет верным.

Пример 2.

Величины X_1, X_2, X_3 имеют совместное нормальное распредлеение с математическим ожиданием $\mathbb{E}(X) = (1, 2, -3)'$ и ковариационной матрицей

$$\mathrm{Var}(X) = \begin{pmatrix} 9 & 2 & -1 \\ 2 & 16 & 1 \\ -1 & 1 & 4 \end{pmatrix}$$

Найдите условную корреляцию $\mathrm{Corr}(X_1, X_2|X_3)$ и частную корреляцию $\mathrm{pCorr}_{X_3}(X_1, X_2)$.

Решение.

...

В данном примере частная и условная корреляция совпали. Это одно из приятных свойств многомерного нормального распределения:

Теорема 4. *Если величины X, Y и Z имеют совместное нормальное распределение, то частная и условная корреляции совпадают.*

Пример 3. AR(1) процесс

7 Выборочные характеристики

В теории обычную корреляцию и частную корреляцию можно посчитать, если известен закон распределения случайных величин. На практике закон распределения не известен, однако доступны наблюдения. Как по имеющимся наблюдениям оценить неизвестные корреляции?

Несколько способов оценки корреляции

Здесь пара картинок: википедийная с корреляциями и два ряда случайного блуждания/тренда

Несколько способов оценки частной корреляции

Доказательство от принцессы

Борис Демешев*

20 сентября 2015 г.

Аннотация

Доказательство от принцессы — частный случай доказательства от противного

Ключевые слова: доказательство от противного, принцесса, доказательство.

1 Доказательство от противного

Как устроено классическое доказательство от противного? Берём кого-нибудь противного, и пусть он доказывает. Допустим нам нужно доказать, что утверждение A верно. Мы, наоборот, предполагаем, что A неверно. Далее каким-нибудь образом приходим к противоречию и, таким образом, получаем вывод, что наше допущение A было ложно.

Довольно часто доказательство от противного используется для того, чтобы доказать, что какой-нибудь объект X не существует. В этом случае очень удобно использовать предлагаемое доказательство от принцессы. Мы представляем себе принцессу, которая замуж не хочет, а по традиции должна объявить конкурс для претендентов руку и сердце. И она объявляет: «Тот, кто принесёт мне X, сможет на мне жениться!». А дальше остаётся объяснить, как она будет аргументированно отказывать каждому претенденту.

2 Пара примеров

Классический пример доказательства от принцессы — доказательство того, что максимальное простое число не существует. Принцесса объявляет: «Тот, кто принесёт мне самое большое просто число во Вселенной, получит меня в жёны!». И к примеру приходит принц и приносит ей p_n. А она ему в ответ: «Не пойду я за тебя замуж, ведь простое число $p_1 \cdot p_2 \cdot p_3 \cdot \ldots \cdot p_n + 1$ больше чем ты принёс!» Так принцесса отказывает всем ухажёрам, а, следовательно, наибольшего простого числа не существует.

Идея доказательства от принцессы возникла так. Я иногда веду вводный курс стохастического анализа для экономистов. Если требуется и позволяет время, то рассказываю про мощности множеств и, в частности, про то, что множество последовательностей из 0 и 1 несчётное. И в нём есть один тонкий момент. Если

*НИУ ВШЭ, Москва.

проводить доказательство в общем виде с произвольными буквами, то оно слишком тяжеловесно. Если проводить на конкретном примере, то возникает вопрос, а почему это доказательство. И принцесса замечательно решает проблему доказательства на частном примере!

Принцесса объявляет: «Тот, кто занумерует натуральными числами все бесконечные последовательности из 0 и 1, получит меня в жёны». И, к примеру, приходит принц датский и говорит: «Я занумеровал!» И предъявляет листочек, на котором все последовательности занумерованы:

1. 000000000...
2. 011001010...
3. 101000000...
4. 010011010...

   ...

Как принцессе отказать принцу датскому? Она выбирает диагональные элементы этих последовательностей 0110... Затем меняет 0 на 1, а 1 на 0, получая 1001... И спрашивает принца датского: «А последовательность 1001... у Вас под каким номером?» И принц датский начинает перебирать. Под первым номером не может идти, так как первой цифрой отличается, под вторым номером не может идти, так как вторым номером отличается... И принц датский трагично вынужден признать, что эту последовательность он забыл занумеровать. И подобным образом принцесса сможет отказать всем претендентам, а значит множество последовательностей несчётно.

ШАД и линал

Артём Филатов*

23 сентября 2015 г.

Аннотация

Ключевые слова: ШАД, линейная алгебра.

1 Кратко про шад

В 2007 году компания Яндекс основала в своих стенах Школу Анализа Данных. Школа была создана для подготовки специалистов в области анализа больших данных, машинного обучения и других смежных дисциплин. Ежегодно в апреле начинаются экзамены, которые проходят в три этапа: онлайн – тест, письменный экзамен и собеседование. Письменный экзамен включает в себя задачи по теории вероятности, алгоритмам, линейной алгебре, математическому анализу и комбинаторике. Предлагаю вам разбор нескольких занимательных задач по линейной алгебре из письменных экзаменов прошлых лет.

2 Задачи по линейной алгебре из шАДовских экзаменов

Задача №1

Дана матрица A размера $n \times n$, где $a_{i,j} = (i-j)^2, i, j = 1, \ldots, n$. Найдите ранг матрицы A.

Решение:

Посмортим, как выглядит наша матрица.

$$A = \begin{pmatrix} 0 & 1 & 4 & \cdots & (n-2)^2 & (n-1)^2 \\ 1 & 0 & 1 & 4 & \cdots & (n-2)^2 \\ 4 & 1 & 0 & 1 & \cdots & (n-3)^2 \\ 9 & 4 & 1 & 0 & \cdots & (n-4)^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ (n-1)^2 & (n-2)^2 & (n-3)^2 & (n-4)^2 & \cdots & 0 \end{pmatrix}$$

*НИУ ВШЭ, Москва.

Из условия каждый элемент матрицы A равен $(i-j)^2 = i^2-2ij+j^2$. Но у матрицы из элементов i^2 ранг 1, у матрицы из элементов j^2 ранг тоже единица. Посмотрим на матрицу, образованную ij:

$$\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ 2 & 4 & 6 & \cdots & 2n \\ 3 & 6 & 9 & \cdots & 3n \\ \cdots & \cdots & \cdots & \cdots & \end{pmatrix}$$

Ее ранг также не превосходит 1. Нам известно, что $\operatorname{rank}(A+B) \le \operatorname{rank}(A)+\operatorname{rank}(B)$, следовательно $\operatorname{rank}(A) \le 3$. Но можно показать, что у нас есть ненулевые миноры 3 порядка, следовательно $\operatorname{rank}(A) = 3$.

Задача №2

Дана матрица из нулей и единиц, причем для каждой строки матрицы верно следующее: если в строке есть единицы, то они все идут подряд. Докажите, что определитель такой матрицы равен 0 или ± 1.

Решение:

Посмротрим на то, как выглядит одна из наших матриц:

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}$$

Переставим строки так, чтобы образовать некое подобие ступенчатой матрицы.

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$

Что произойдет с определителем? Он либо не изменился, либо изменил знак, так как перестановка строк меняет знак определителя на противоположный. Теперь сделаем следующее: если позиции первых единиц у строк совпали, то вычтем из той в которой больше единиц, ту в которой меньше единиц. На определитель данное преобразование никак не влияет.

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$

Переставляя строки и повторяя данную процедуру, мы получим ступенчатую матрицу, которая будет либо вырождена, либо иметь единицы на диагонали. А так как в такой матрице $\det(A) = \prod_{i=1} a_{i,i} = 1$, то детерминант исходной матрицы равен либо 0, либо ± 1.

Задача №3

Опишите все невырожденные вещественные матрицы A, для которых все элементы матриц A и A^{-1} неотрицательны.

Решение:

Пусть исходная невырожденная матрица А заполнена некотрыми элементами $a_{i,j}$, а обратная к ней A^{-1} элементами $b_{i,j}$. Как известно, $AA^{-1} = E$. Значит, произведение первой строки на первый столбец должно дать 1:

$$a_{1,1} \cdot b_{1,1} + a_{1,2} \cdot b_{2,1} + \dots + a_{1,n} \cdot b_{n,1} = 1$$

Но произведение первой строки на все остальные столбцы должно дать 0, также нам известно, что все элементы матриц неотрицательны, значит если $a_{1,i} \neq 0$, то $b_{i,j} = 0, j = 2 \dots n$. Это должно быть выполнено для всех $a_{i,j}$. Формально:

$$a_{i,j} \neq 0 \Rightarrow b_{j,z} = 0, z = 1 \dots (i-1), (i+1) \dots n$$

Докажем, что нет такой матрицы A с двумя и более положительными элементами в одном столбце:

Зафиксируем столбец j. Предположим мы встретили первый ненулевой элемент, тогда все кроме одного элементы в j строке матрицы A^{-1} равны 0. Предположим, что мы встретили второй положительны элемент, тогда он занулит все элементы кроме одного, включая тот, который сы не занулили в первый раз. Следовательно, мы получили, что $b_{j,z} = 0$ для всех $z = 1 \dots n$. Но это невозможно, так как это означало бы, что все алгебраические дополнения в некой строке матрицы A равны 0 ($b_{i,j} = \frac{Alg_{i,j}}{\det(A)}$), а следовательно и определитель.

Из всего сказанного следует, что единственно законной матрицей A будет такая матрица, в столбцах которых по одному положительному элементу. Элементарными преобразованиями такая матрица приводится к диагональному виду. Мы показали, что все элементы обратной матрицы зануляться, кроме тех, которые образуют 1 в произведении с ненулевыми элементами матрицы A, следовательно обратная матрица будет иметь аналогичный вид.

Задача №4

Имеется некоторый ненулевой вектор – столбец v. Найти все собственные значения матрицы $v \cdot v^T$.

Решение:

Первым делом необходимо понять сколько собственных значений нужно найти. Нам известно, что $\operatorname{rank}(A) \cdot \operatorname{rank}(B) \leq \min \operatorname{rank}(A), \operatorname{rank}(B)$. Следовательно итоговая матрица будет иметь ранг равный единице. Ранг при замене базиса не изменяется, тогда мы можем перейти к диагональному виду матрицы с базисом из собственных векторов, где на диагонали будет лишь одно собственное значение. Собственные значения также не изменяются при замене базиса. Осталось его найти!

Для того, чтобы найти собственное значение (можно догадаться чему оно равно) воспользуемся ещё одним интересным свойством. Оказывается, что след матрицы (сумма диагональных элементов) равен сумме собственных значений матрицы с учётом кратности. Легко увидеть, что сумма диагональных элементов это скалярное произведение вектора на самого себя.

Следовательно, мы имеем собственное значение $v^T \cdot v$, и нулевое собственное значение кратности $n - 1$.

Случайная перестановка (рабочее название)

Борис Демешев*

20 сентября 2015 г.

Аннотация

Случайная перестановка

Ключевые слова: задача, случайная перестановка, киллер.

1 Сумасшедшая старушка

В самолете 100 мест и все билеты проданы. Первой в очереди на посадку стоит Сумасшедшая Старушка. Сумасшедшая Старушка очень переживает, что ей не хватит места, врывается в самолёт и несмотря на номер по билету садиться на случайно выбираемое место. Каждый оставшийся пассажир садится на своё место, если оно свободно, и на случайное выбираемое место, если его место уже кем-то занято.

1. Какова вероятность того, что последний пассажир сядет на своё место?
2. Чему примерно равно среднее количество пассажиров севших на свои места?

2 Судьба Дон-Жуана

У Дон-Жуана n знакомых девушек, и их всех зовут по-разному. Он пишет им n писем, но по рассеянности раскладывает их в конверты наугад. Случайная величина X обозначает количество девушек, получивших письма, адресованные лично им.

1. Найдите $\mathbb{E}(X)$, $\operatorname{Var}(X)$
2. Какова при большом n вероятность того, что хотя бы одна девушка получит письмо, адресованное ей?

3 Киллер

Правила игры «Киллер» просты. Игроки пишут на бумажках, как их зовут, и кладут бумажки в шляпу. Каждый тянет из шляпы имя своей первой жертвы. Если первой жертвой игрока является он сам, то он совершает «самоубийство» и дальше

*НИУ ВШЭ, Москва.

не играет[1]. Чтобы убить жертву, надо остаться с ней наедине и сказать: «Ты убит!». Убийца забирает себе все бумажки, набранные убитым, и начинается охотиться за тем, за кем охотился убитый. Побеждает тот, кто наберёт больше всех бумажек к концу игры. Заметим, что в «Киллере» каждый игрок оказывается втянут в одну из нескольких цепочек.

В «Киллера» играют 30 человек, из них 20 девушек.

1. Какова вероятность того, что в цепочке, начинающейся с Маши Сидоровой ровно 5 человек?
2. Какова вероятность того, что в цепочке, начинающейся с Маши Сидоровой ровно 5 девушек?
3. Какова вероятность того, что все девушки попадают в одну цепочку убийц и жертв?
4. Какова вероятность того, что все игроки попадают в общую цепочку?
5. Сколько в среднем цепочек в «Киллере»?
6. Сколько в среднем «самоубийц»?

4 Ключи и копилки

На столе стоят n свиней-копилок. Достать содержимое копилки можно двумя способами: либо разбить копилку, либо открыть дно специальным ключиком. К каждой копилке подходит единственный ключ. Мы раскладываем ключи по копилкам наугад, один ключ в одну копилку. Затем разбиваем k копилок и получаем хранящиеся в них ключи. Далее мы будем копилки только открывать ключами.

1. Какова вероятность того, что мы сможем достать все ключи?
2. Какая доля ключей в среднем будет найдена?

5 Задача о макаронах

В тарелке запутавшись лежат $n >> 0$ макаронин. Я по очереди связываю попарно все торчащие концы макаронин.

1. Какова примерно вероятность того, что я свяжу все макаронины в одно большое кольцо?
2. Сколько в среднем колец образуется?
3. Каково среднее число колец длиной в одну макаронину?

6 Задача о 100 заключенных

У ста узников тюрьмы есть последний шанс на спасение. В комнате стоит шкаф в котором сто занумерованных ящичков. Палач кладёт в каждый ящичек бумажку с номером ровно одного из заключенных в случайном порядке и задвигает все ящички. Узники заходят в комнату один за другим. Каждый узник может открыть любые 50 ящичков. После каждого узника все ящички задвигаются в исходное положение.

[1] В некоторых вариантах правил, если игрок вытянул из шляпы своё имя, то он должен вытянуть другую бумажку.

Если каждый узник находит свой номер, то все узники будут помилованы. Если хотя бы один из узников не найдёт свой номер, то все будут казнены.

Узники могут предварительно договориться о стратегии.

1. Какова оптимальная стратегия?
2. Какую вероятность выигрыша она обеспечивает?

7 что-то про детерминант?