1 Два способа задать геометрию

Множество случайных величин - это векторное (линейное) пространство. Если сложить две случайных величины, то получится случайная величина, если умножить случайную величину на число получится случайная величина. Поэтому случайные величины - это векторы.

Чтобы задать геометрию достаточно определить скалярное произведение, то есть действие $< X, Y >$, которое любой паре случайных величин ставит в соответствие число. При этом должны выполняться свойства:

1. $< X, Y >=< Y, X >$
2. $< X + Y, Z >=< X, Z > + < Y, Z >$
3. $< X, X >\geqslant 0$
4. $< X, X >= 0$ только если $X = 0$.

Почему скалярное произведение определяет геометрию? Геометрия - это же про длины, углы и расстояния! Все это восстанавливается из скалярного произведения по формулам 9-го класса. Длину любого вектора теперь можно найти по формуле $||X|| = \sqrt{< X, X >}$. Для того, чтобы найти угол между векторами достаточно знать косинус этого угла. А косинус определяется как: $cos(X, Y) = \frac{<X,Y>}{||X||||Y||}$. Расстояние определяется как длина разности: $d(X, Y) = ||X - Y||$.

Мы используем двойные палочки для длины чтобы отличать ее от модуля: модуль случайной величины - это случайная величина, а длина - это константа.

Можно предложить много разных геометрий (или, что то же самое, скалярных произведений) в пространстве случайных величин, но интересными, пожалуй, являются две: $< X, Y >= \mathbb{E}(XY)$ и $< X, Y >= \mathrm{Cov}(X, Y)$.

Геометрия позволит «увидеть» некоторые понятия и теоремы. Чаще всего мы будет сталкиваться с теоремой Пифагора, настолько часто, что можно быть уверенным: если что-то неотрицательное равно сумме двух неотрицательных частей, то это теорема Пифагора и можно ее проиллюстрировать.

2 Геометрия ожидаемого произведения

Пусть скалярное произведение задано $< X, Y >= \mathbb{E}(XY)$.

Легко убедиться, что первые три требования к скалярному произведению выполнены:

1. $\mathbb{E}(XY) = \mathbb{E}(YX)$
2. $\mathbb{E}((X + Y)Z) = \mathbb{E}(XZ) + \mathbb{E}(YZ)$
3. $\mathbb{E}(X^2) \geqslant 0$

Четвертое требование выполнено не совсем полностью. Оказывается $\mathbb{E}(X^2)$ может быть равно нулю, даже если X не всегда ноль. Например, пусть Y равномерно на $[0; 1]$, а X равен 1, если $Y = 0.5$ и 0 иначе. В данном примере X может равняться одному, но $\mathbb{E}(X^2) = 0$. Вызвано это тем, что $P(X = 0) = 1$. То есть в этой геометрии нулевой считается случайная величина, которая равна нулю с вероятностью один.

Если вероятность события A равна 1, то говорят, что A происходит почти наверное. В геометрии ожидаемого произведения случайные величины почти наверное равные нулю не отличимы от настоящего нуля. И, следовательно, если $X = Y$ почти наверное, то в этой геометрии X не отличим от Y. Действительно, в этом случае $X - Y = 0$ почти наверное, то есть $X - Y$ не отличима от нуля.

Что нам дает введение геометрии?

Теперь вполне серьезно можно говорить о длине случайное величины $||X|| = \sqrt{\mathbb{E}(X^2)}$ или об угле между двумя случайными величинами $\angle(X, Y) = arccos(\frac{\mathbb{E}(XY)}{\sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}})$. А что из этого?

Автоматически возникает понятие перпендикулярных (ортогональных) векторов: случайные величины перпендикулярны, если угол между ними равен $\frac{\pi}{2}$. Или, если косинус угла между ними равен нулю. Или, $\mathbb{E}(XY) = 0$.

Теорему Пифагора никто не отменял: если $X \perp Y$, то $||X-Y||^2 = ||X||^2 + ||Y||^2$.

Доказательство: $\mathbb{E}((X-Y)^2) = \mathbb{E}(X^2) + \mathbb{E}(Y^2) - 2\mathbb{E}(XY) = \mathbb{E}(X^2) + \mathbb{E}(Y^2)$.

Заметим пару интересных фактов в нашей геометрии:

Длина любой константы равна ее модулю: $||c|| = \sqrt{\mathbb{E}(c^2)} = \sqrt{c^2} = |c|$, или $||c||^2 = c^2$.

Если $\mathbb{E}(Y) = 0$, то случайная величина Y перпендикулярна любой константе: $< Y, c >= \mathbb{E}(Yc) = c\mathbb{E}(Y) = 0$.

Рисунок 1. числовая прямая, ей перпендикулярная величина Y и неперпендикулярная X

Применим теорему Пифагора чтобы увидеть дисперсию...

Пусть X - произвольная случайная величина. У случайной величины $Y = X - \mathbb{E}(X)$ матожидание равно нулю, $\mathbb{E}(Y) = \mathbb{E}(X) - \mathbb{E}(X) = 0$, поэтому Y перпендикулярна любой константе, в частности, $Y \perp \mathbb{E}(X)$. Заметим кстати, что $||X - \mathbb{E}(X)||^2 = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathrm{Var}(X)$.

Применяя теорему Пифагора к $X - \mathbb{E}(X) \perp \mathbb{E}(X)$ получаем: $||X||^2 = ||X - \mathbb{E}(X)||^2 + ||\mathbb{E}(X)||^2$

Переходя к ожиданиям, получаем $\mathbb{E}(X^2) = \mathrm{Var}(X) + (\mathbb{E}(X))^2$.

Рисунок 2. числовая прямая, ей неперпендикулярная X, подписи $\mathbb{E}(X)^2$...

Из рисунка видно, что $\mathbb{E}(X)$ это проекция случайной величины X на множество констант! Это действительно так в нашей геометрии:

Во-первых, $X - \mathbb{E}(X) \perp \mathbb{E}(X)$.

Во-вторых, если взять любую другую константу $c \neq \mathbb{E}(X)$, то расстояние от X до этой константы c будет больше, чем до константы $\mathbb{E}(X)$: $||X - c|| > ||X - \mathbb{E}(X)||$. Доказательство: Рассмотрим функцию $\mathbb{E}((X-c)^2) = \mathbb{E}(X^2) + c^2 - 2c\mathbb{E}(X)$. Относительно c это парабола с ветвями вверх и вершиной при $c = \mathbb{E}(X)$. Значит наименьшое значение функции равно $||X - \mathbb{E}(X)||$.

3 Связь со школьной геометрией

4 Условное ожидание - это проекция!

5 Геометрия ковариации

Пусть скалярное произведение задано $< X, Y >= \mathrm{Cov}(X, Y)$.

Все требования к скалярному произведению кроме ... выполнены.

Требование ... выполнено с оговорками.

Эта геометрия наглядна тем, что некоррелированные случайные величины в ней перпендикулярны. Если $Corr(X, Y) = 0$, то $\mathrm{Cov}(X, Y) = 0$ и, следовательно, X и Y ортогональны.

Напомним, что независимость означает некоррелированность любых[1] функций $f(X)$ и $g(Y)$.

6 Частная корреляция

В анализе временных рядов при изучении процессов ARMA используется понятие частной корреляции. При этом указывается некий шаманский способ ее подсчета (как правило это система уравнений Юла-Воркера) и изредка - ее интуитивная интерпретация. Мы же беремся рассказать что это такое на самом деле в рамках геометрии ковариации! Станет ясна связь между формулой расчета и интуивной интерпретацией!

Рассмотрим три случайных величины X, Y и Z. Обычная корреляция $Corr(X, Y)$ - это косинус угла между X и Y, $cos(X, Y)$. Что же такое частная корреляция X Y при фиксированном (????) Z, $Corr(X, Y|Z)$ (????).

Очень просто! Изначально X коррелировано с Z и Y коррелировано с Z. Возьмем и «очистим» X и Y от воздействия Z, то есть найдем самые похожие на них величины $\hat{X}$ и $\hat{Y}$ не

[1] Любых борелевских функций (для знакомых с теорией меры)

коррелированные с Z. От случайной величины $\hat{X}$ требуется чтобы она была некоррелирована с Z и максимально похожа на X, то есть чтобы квадрат расстояния $||X - \hat{X}||^2 = \mathrm{Var}(X - \hat{X})$ был минимальным. Геометрически это означает следующее: есть $Z^{\perp}$ - множество случайных величин, некоррелированных с Z (ортогональных Z). Просто спроецируем X и Y на это множество $Z^{\perp}$. Так частная корреляция между X и Y при фиксированном Z это просто обычная корреляция между $\hat{X}$ и $\hat{Y}$, $Corr(X, Y|Z) = Corr(\hat{X}, \hat{Y})$ или косинус угла между проекциями X и Y на плоскость $Z^{\perp}$.

Давайте попробуем посчитать на простом примере, не связанном с временными рядами. Пусть X, Y, Z

Из этого определения следует формула расчета, предлагаемая процедурой Юла-Воркера.

Чтобы изложения было законченным - проиллюстрируем интуитивную интерпретацию

то есть спроецируем величины X и Y на множество случайных величин