

ВЗАИМНАЯ ИНФОРМАЦИЯ ДВУХ СЛУЧАЙНЫХ ПОЛЕЙ

Найдено выражение для определения взаимной информации двух случайных статистически однородных гауссовских полей.

Различные виды деятельности человека все в большей мере связаны с применением и обработкой сведений о случайных (по меньшей мере, двумерных) полях. К ним относятся фотографии различного типа, результаты измерений физических полей в гидрометеорологии и океанологии, данные, получаемые при сейсморазведке полезных ископаемых и др. Часто различные поля оказываются статистически связанными друг с другом. Это является основой извлечения сведений об одном поле из результатов измерения другого поля, позволяет устранить избыточную информацию при архивировании данных о них и т.д.

Для оценки эффективности этих процедур необходимо иметь аналитическое выражение для взаимной информации статистически связанных полей. В литературе имеются формулы для взаимной информации двух случайных величин, двух систем случайных величин, пары случайных процессов, двух групп случайных процессов [1], однако отсутствуют выражения, позволяющие найти удельную взаимную информацию двух случайных полей. Данная работа имеет целью заполнить этот пробел.

Найдем аналитическое выражение для средней удельной взаимной информации двух случайных двумерных скалярных статистически однородных стационарно связанных полей, распределенных по нормальному закону.

Для этого вначале обратимся к двум множествам центрированных нормально распределенных случайных величин — векторам $\xi^T = (\xi_1, \xi_2, \dots, \xi_m)$ и $\eta^T = (\eta_1, \eta_2, \dots, \eta_m)$ одинакового размера m . Здесь знак T означает транспонирование. Будем считать, что эти множества описываются следующими квадратными матрицами размера $m \times m$:

— матрица корреляционных моментов между составляющими вектора ξ^T

$$B_{\xi} = \begin{bmatrix} M[\xi_1 \xi_1] & M[\xi_1 \xi_2] & \dots & M[\xi_1 \xi_m] \\ M[\xi_2 \xi_1] & M[\xi_2 \xi_2] & \dots & M[\xi_2 \xi_m] \\ \dots & \dots & \dots & \dots \\ M[\xi_m \xi_1] & M[\xi_m \xi_2] & \dots & M[\xi_m \xi_m] \end{bmatrix},$$

— матрица корреляционных моментов между составляющими вектора η^T

$$B_{\eta} = \begin{bmatrix} M[\eta_1 \eta_1] & M[\eta_1 \eta_2] & \dots & M[\eta_1 \eta_m] \\ M[\eta_2 \eta_1] & M[\eta_2 \eta_2] & \dots & M[\eta_2 \eta_m] \\ \dots & \dots & \dots & \dots \\ M[\eta_m \eta_1] & M[\eta_m \eta_2] & \dots & M[\eta_m \eta_m] \end{bmatrix}$$

и матрица взаимных корреляционных моментов между составляющими векторов ξ^T и η^T

$$B_{\xi\eta} = \begin{bmatrix} M[\xi_1 \eta_1] & M[\xi_1 \eta_2] & \dots & M[\xi_1 \eta_m] \\ M[\xi_2 \eta_1] & M[\xi_2 \eta_2] & \dots & M[\xi_2 \eta_m] \\ \dots & \dots & \dots & \dots \\ M[\xi_m \eta_1] & M[\xi_m \eta_2] & \dots & M[\xi_m \eta_m] \end{bmatrix}.$$

В работе [1] показано, что выражение для средней взаимной информации о множестве действительных величин ξ^T , содержащейся в множестве также действительных величин η^T (или, что то же, о множестве величин ξ^T , содержащейся в множестве η^T), дается формулой

$$I(\xi^T, \eta^T) = -\frac{1}{2} \log \det (E - B_{\xi\eta} B_{\eta}^{-1} B_{\xi\eta}^T B_{\xi}^{-1}),$$

где E — единичная матрица соответствующего размера.

Если же величины, из которых состоят множества ξ^T и η^T , комплексные, то транспонирование матриц необходимо дополнить комплексным сопряжением, обозначаемым знаком $*$. В этом случае [2]

$$I(\xi^T, \eta^T) = -\frac{1}{2} \log \det (E - B_{\xi\eta} B_{\eta}^{-1} B_{\xi\eta}^* B_{\xi}^{-1}). \quad (1)$$

Теперь вернемся к решению поставленной задачи.

Выделим в двумерном поле $X(x, y)$ квадрат размера $L \times L$, считая, что $-L/2 \leq x \leq L/2$ и $-L/2 \leq y \leq L/2$. Рассмотрим на нем случайное скалярное комплексное поле вида

$$X(x, y) = \sum_{k=-n}^n \sum_{l=-n}^n \Phi_{kl} \exp[j(\alpha_k x + \beta_l y)], \quad (2)$$

где обозначены как

$$\alpha_k = 2\pi k/L, \quad \beta_l = 2\pi l/L \quad (3)$$

волновые числа соответственно вдоль оси абсцисс и оси ординат. Допустим, что Φ_{kl} — некоррелированные комплексные случайные величины с нулевыми средними значениями, т.е.

$$M[\Phi_{kl}] = 0 \text{ при всех } k \text{ и } l \quad (4)$$

и

$$M[\Phi_{kl} \Phi_{rs}^*] = \begin{cases} 0, & \text{если } r \neq k \text{ или } s \neq l, \\ M[|\Phi_{kl}|^2] = f_{kl} & \text{при } r = k \text{ и } s = l, \end{cases} \quad (5)$$

где $-n \leq k, l, r, s \leq n$. Вещественные положительные числа f_{kl} представляют собой мощность соответствующих гармонических составляющих поля $X(x, y)$.

Очевидно, что математическое ожидание такого поля $M[X(x, y)] = 0$. Двумерная корреляционная функция этого поля внутри выбранного квадрата есть

$$\begin{aligned} B_X(x_1, x_2; y_1, y_2) &= M[X(x_1, y_1) X^*(x_2, y_2)] = \\ &= \sum_{k=-n}^n \sum_{l=-n}^n \sum_{r=-n}^n \sum_{s=-n}^n M[\Phi_{kl} \Phi_{rs}^*] \exp[j(\alpha_k x_1 - \alpha_r x_2) + j(\beta_l y_1 - \beta_s y_2)]. \end{aligned}$$

В силу равенств (5) это выражение упрощается:

$$B_X(x_1, x_2; y_1, y_2) = \sum_{k=-n}^n \sum_{l=-n}^n f_{kl} \exp[j\alpha_k (x_1 - x_2) + j\beta_l (y_1 - y_2)]. \quad (6)$$

Корреляционная функция (6) зависит только от сдвигов $\rho_1 = x_1 - x_2$ и $\rho_2 = y_1 - y_2$ соответственно вдоль оси абсцисс и оси ординат и может быть записана как $B_X(\rho_1, \rho_2)$. Следовательно, описываемое моделью (4), (5) случайное поле в пределах рассматриваемого квадрата является комплексным центрированным статистически однородным (по крайней мере в широком смысле) с дискретным спектром. За пределами квадрата оно периодически повторяется с периодом L как вдоль оси абсцисс, так и вдоль оси ординат. Соответственно периодически повторяется и его корреляционная функция. Средняя мощность этого поля (т.е. его дисперсия)

$$B_X(0, 0) = \sum_{k=-n}^n \sum_{l=-n}^n f_{kl}$$

согласно равенству (6) складывается из мощности (дисперсий) f_{kl} его отдельных составляющих.

Предположим далее, что в рассматриваемом квадрате кроме поля $X(x, y)$ имеется другое поле $Y(x, y)$ такого же типа

$$Y(x, y) = \sum_{k=-n}^n \sum_{l=-n}^n \Psi_{kl} \exp[j(\alpha_k x + \beta_l y)], \quad (7)$$

отличающееся от первого случайными коэффициентами Ψ_{kl} . Для него выполняются равенства, аналогичные (4), (5):

$$\begin{aligned} M[\Psi_{kl}] &= 0 \text{ при всех } k \text{ и } l, \\ M[\Psi_{kl} \Psi_{rs}^*] &= \begin{cases} 0, & \text{если } r \neq k \text{ или } s \neq l, \\ M[|\Psi_{kl}|^2] = g_{kl} & \text{при } r = k \text{ и } s = l, \end{cases} \end{aligned} \quad (8)$$

где g_{kl} — вещественные положительные числа, представляющие собой мощности соответствующих спектральных составляющих поля (7).

Будем считать, что поля $X(x, y)$ и $Y(x, y)$ стационарно связаны. Это означает, что их взаимная корреляционная функция зависит от сдвига. Найдем условия, при которых это имеет место. Взаимная корреляционная функция определяется выражением

$$B_{XY}(x_1, x_2; y_1, y_2) = M[X(x_1, y_1) Y^*(x_2, y_2)].$$

Подставим в него формулы (2) и (7). Тогда получим

$$B_{XY}(x_1, x_2; y_1, y_2) = \sum_{k=-n}^n \sum_{l=-n}^n \sum_{r=-n}^n \sum_{s=-n}^n M[\Phi_{kl} \Psi_{rs}^*] \exp[j(\alpha_k x_1 - \alpha_r x_2) + j(\beta_l y_1 - \beta_s y_2)].$$

Для того чтобы эта функция зависела только от сдвигов, необходимо и достаточно, чтобы выполнялись условия

$$M[\Phi_{kl} \Psi_{rs}^*] = \begin{cases} 0, & \text{если } r \neq k \text{ или } s \neq l, \\ M[\Phi_{kl} \Psi_{kl}^*] = h_{kl} & \text{при } r = k \text{ и } s = l, \end{cases} \quad (9)$$

где h_{kl} — число, в общем случае комплексное, характеризующее корреляцию совпадающих номеров спектральных составляющих полей $X(x, y)$ и $Y(x, y)$. При различных номерах эти составляющие не коррелируют друг с другом. Нетрудно показать, что

$$|h_{kl}|^2 \leq f_{kl} g_{kl}. \quad (10)$$

Итак, каждое из статистически однородных случайных полей $X(x, y)$ и $Y(x, y)$ определяется множеством $(2n+1)^2$ случайных комплексных величин Φ_{kl} и Ψ_{kl} , удовлетворяющих условиям (5), (8) и (9). Именно они содержат как собственную, так и взаимную информацию о полях $X(x, y)$ и $Y(x, y)$. Как известно, неслучайные функции (здесь это экспоненциальные функции в представлениях (2) и (7)) информации не несут.

Обозначим через ξ_k^T вектор, компонентами которого являются случайные величины Φ_{kl} , где $-n \leq l \leq n$, т.е. $\xi_k^T = (\Phi_{k,-n}, \Phi_{k,-n+1}, \dots, \Phi_{k,-1}, \Phi_{k,0}, \Phi_{k,1}, \dots, \Phi_{k,n-1}, \Phi_{k,n})$. Составляющие этого вектора согласно (5) не коррелированы. Поэтому соответствующая им корреляционная матрица диагональна:

$$B_k(\xi) = \begin{bmatrix} f_{k,-n} & 0 & 0 & \dots & 0 \\ 0 & f_{k,-n+1} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & f_{k,n} \end{bmatrix}. \quad (11)$$

В представлении (2) имеем $2n+1$ таких векторов, так как $-n \leq k \leq n$. В совокупности они дают составной вектор

$$\xi^T = (\xi_{-n}^T, \xi_{-n+1}^T, \dots, \xi_{-1}^T, \xi_0^T, \xi_1^T, \dots, \xi_{n-1}^T, \xi_n^T). \quad (12)$$

Его корреляционная матрица представляет собой диагональную блочную матрицу, состоящую из $(2n+1)$ блоков вида (11) разных номеров:

$$B_\xi = \begin{bmatrix} B_{-n}(\xi) & 0 & 0 & \dots & 0 \\ 0 & B_{-n+1}(\xi) & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & B_n(\xi) \end{bmatrix}. \quad (13)$$

Очевидно, обратная ей матрица есть

$$B_\xi^{-1} = \begin{bmatrix} B_{-n}^{-1}(\xi) & 0 & 0 & \dots & 0 \\ 0 & B_{-n+1}^{-1}(\xi) & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & B_n^{-1}(\xi) \end{bmatrix}.$$

Входящие в нее блоки согласно (11) диагональны:

$$B_k^{-1}(\xi) = \begin{bmatrix} 1/f_{k,-n} & 0 & 0 & \dots & 0 \\ 0 & 1/f_{k,-n+1} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1/f_{k,n} \end{bmatrix}.$$

Аналогичным образом обозначим η_k^T вектор с компонентами — случайными величинами Ψ_{kl} , где $-n \leq l \leq n$: $\eta_k^T = (\Psi_{k,-n}, \Psi_{k,-n+1}, \dots, \Psi_{k,-1}, \Psi_{k,0}, \Psi_{k,1}, \dots, \Psi_{k,n-1}, \Psi_{k,n})$. Его корреляционная матрица также диагональна:

$$B_k(\vec{\eta}) = \begin{bmatrix} g_{k,-n} & 0 & 0 & \dots & 0 \\ 0 & g_{k,-n+1} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & g_{kn} \end{bmatrix}. \quad (14)$$

Из векторов $\vec{\eta}_k^T$ образуем составной вектор

$$\vec{\eta}^T = (\vec{\eta}_{-n}^T, \vec{\eta}_{-n+1}^T, \dots, \vec{\eta}_{-1}^T, \vec{\eta}_0^T, \vec{\eta}_1^T, \dots, \vec{\eta}_{n-1}^T, \vec{\eta}_n^T). \quad (15)$$

Подобно формуле (13), его матрица имеет диагональный блочный вид

$$B_{\vec{\eta}} = \begin{bmatrix} B_{-n}(\vec{\eta}) & 0 & 0 & \dots & 0 \\ 0 & B_{-n+1}(\vec{\eta}) & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & B_n(\vec{\eta}) \end{bmatrix}.$$

Обратная ей матрица есть

$$B_{\vec{\eta}}^{-1} = \begin{bmatrix} B_{-n}^{-1}(\vec{\eta}) & 0 & 0 & \dots & 0 \\ 0 & B_{-n+1}^{-1}(\vec{\eta}) & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & B_n^{-1}(\vec{\eta}) \end{bmatrix}.$$

Входящие в нее блоки согласно (14) диагональны:

$$B_k^{-1}(\vec{\eta}) = \begin{bmatrix} 1/g_{k,-n} & 0 & 0 & \dots & 0 \\ 0 & 1/g_{k,-n+1} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1/g_{kn} \end{bmatrix}.$$

В силу условий $f_{kl}, g_{kl} > 0$ матрицы $B_k(\vec{\xi})$, $B_k(\vec{\eta})$, $B_{\vec{\xi}}^{-1}$, $B_{\vec{\eta}}^{-1}$ — неособенные.

Рассмотрим взаимную корреляцию векторов (12) и (15). Согласно соотношению (9), коррелируют между собой только их одноименные компоненты. Это означает, что матрица их взаимной корреляции также является диагональной блочной:

$$B_{\vec{\xi}\vec{\eta}} = \begin{bmatrix} B_{-n}(\vec{\xi}, \vec{\eta}) & 0 & 0 & \dots & 0 \\ 0 & B_{-n+1}(\vec{\xi}, \vec{\eta}) & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & B_n(\vec{\xi}, \vec{\eta}) \end{bmatrix}$$

и ее блоки диагональны:

$$B_k(\vec{\xi}, \vec{\eta}) = \begin{bmatrix} h_{k,-n} & 0 & 0 & \dots & 0 \\ 0 & h_{k,-n+1} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & h_{kn} \end{bmatrix}.$$

Решим задачу, поставленную в начале статьи. Для этого вначале найдем величину взаимной информации между векторами $\vec{\xi}^T$ и $\vec{\eta}^T$. С этой целью обратимся к формуле (1) и вычислим входящую в ее аргумент матрицу, которую обозначим как $B(I) = E - B_{\vec{\xi}\vec{\eta}} B_{\vec{\eta}}^{-1} B_{\vec{\xi}}^{T*} B_{\vec{\xi}}^{-1}$. Путем простых вычислений нетрудно убедиться, что она представима в виде

$$B(I) = \begin{bmatrix} B_{-n}(I) & 0 & 0 & \dots & 0 \\ 0 & B_{-n+1}(I) & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & B_n(I) \end{bmatrix},$$

т.е. также состоит из расположенных на главной диагонали блочных диагональных матриц

$$B_k(I) = \begin{bmatrix} 1 - |h_{k,-n}|^2 / (f_{k,-n} g_{k,-n}) & 0 & 0 & \dots & 0 \\ 0 & 1 - |h_{k,-n+1}|^2 / (f_{k,-n+1} g_{k,-n+1}) & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 - |h_{kn}|^2 / (f_{kn} g_{kn}) \end{bmatrix}.$$

Определитель диагональной матрицы $B(I)$ равен произведению ее диагональных элементов, которые сами являются диагональными матрицами и в силу неравенства (10) неотрицательны:

$$\det B(I) = \prod_{k=-n}^n \prod_{l=-n}^n \left(1 - \frac{|h_{kl}|^2}{f_{kl} g_{kl}} \right).$$

Подставим это выражение в формулу (1). При этом получим

$$I(\vec{\xi}, \vec{\eta}) = -\frac{1}{2} \sum_{k=-n}^n \sum_{l=-n}^n \log \left(1 - \frac{|h_{kl}|^2}{f_{kl} g_{kl}} \right). \quad (16)$$

Если число составляющих разложений (2) и (7) бесконечно, то формула (16) принимает вид

$$I(\vec{\xi}, \vec{\eta}) = -\frac{1}{2} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \log \left(1 - \frac{|h_{kl}|^2}{f_{kl} g_{kl}} \right). \quad (17)$$

Введем понятие средней на единицу площади удельной взаимной информации рассматриваемых двумерных полей X и Y с помощью соотношения

$$i_2(X, Y) = I(\vec{\xi}, \vec{\eta}) / L^2, \quad (18)$$

где L — длина стороны введенного выше квадрата. Устремим теперь L к бесконечности. Это ведет к уменьшению интервалов между соседними волновыми числами (3) полей $X(x, y)$ и $Y(x, y)$, равных $\Delta\alpha = \alpha_{k+1} - \alpha_k = 2\pi/L$ и $\Delta\beta = \beta_{k+1} - \beta_k = 2\pi/L$. При этом мощности f_{kl} и g_{kl} их соответствующих составляющих можно представить как

$$f_{kl} = G_{2X}(\alpha, \beta) \Delta\alpha \Delta\beta, \quad g_{kl} = G_{2Y}(\alpha, \beta) \Delta\alpha \Delta\beta, \quad (19)$$

где $G_{2X}(\alpha, \beta)$ и $G_{2Y}(\alpha, \beta)$ — двумерные энергетические спектральные плотности полей X и Y . Величина h_{kl} приобретает вид

$$h_{kl} = G_{2XY}(\alpha, \beta) \Delta\alpha \Delta\beta, \quad (20)$$

где $G_{2XY}(\alpha, \beta)$ — взаимный двумерный спектр этих полей.

Подставим в соотношение (18) формулу (17) с учетом (19) и (20) и осуществим предельный переход при $L \rightarrow \infty$. Тогда получим

$$i_2(X, Y) = -\frac{1}{2(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \log \left[1 - \frac{|G_{2XY}(\alpha, \beta)|^2}{G_{2X}(\alpha, \beta) G_{2Y}(\alpha, \beta)} \right] d\alpha d\beta. \quad (21)$$

Обозначая волновой вектор как $\vec{\alpha} = \alpha \vec{e}_1 + \beta \vec{e}_2$, выражение (21) можно представить в виде

$$i_2(X, Y) = -\frac{1}{2(2\pi)^2} \int \log \left[1 - \frac{|G_{2XY}(\vec{\alpha})|^2}{G_{2X}(\vec{\alpha}) G_{2Y}(\vec{\alpha})} \right] d\vec{\alpha}, \quad (22)$$

где интегрирование производится по всей области существования двумерных спектров. Так как $L \rightarrow \infty$, поля X и Y здесь непериодичны и существуют во всем безграничном пространстве.

Возможно еще одно упрощение формулы (21). В ней подынтегральная функция четна, поэтому двустороннее интегрирование можно заменить односторонним:

$$i_2(X, Y) = -\frac{1}{2\pi^2} \int_0^{\infty} \int_0^{\infty} \log \left[1 - \frac{|G_{2XY}(\alpha, \beta)|^2}{G_{2X}(\alpha, \beta) G_{2Y}(\alpha, \beta)} \right] d\alpha d\beta. \quad (23)$$

Формулы (21) — (23) решают поставленную в начале статьи задачу.

Рассуждая образом, аналогичным изложенному выше, можно показать, что средняя удельная взаимная информация двух статистически однородных стационарно связанных гауссовских скалярных N -мерных полей $X(x_1, x_2, \dots, x_N)$ и $Y(x_1, x_2, \dots, x_N)$, отнесенная к объему соответствующей размерности, дается формулой

$$i_N(X, Y) = -\frac{1}{2(2\pi)^N} \int \log \left[1 - \frac{|G_{NX}(r)G_{NY}(r)|^2}{G_{NX}(r)G_{NY}(r)} \right] d\mathbf{r},$$

где волновой вектор $\mathbf{r}$ N -мерен и интегрирование производится по всей области существования подынтегрального выражения. В частности, при $N=1$ отсюда получим формулу для взаимной удельной информации двух стационарных процессов [1,2].

Полученные соотношения дают возможность ставить и решать теоретические задачи о взаимосвязи случайных составляющих физических полей в различных областях сейсмологии, геофизики, гидрофизики, гидрометеорологии и т.д. Это будет способствовать рациональной организации системы измерений таких полей, их эффективной математической обработке и интерпретации результатов исследований. В этом — перспектива дальнейших исследований в данном направлении.

Библиографический список

1. Стратонович Р.Л. Теория информации / Р.Л. Стратонович. — М.: Сов. радио, 1975. — 424 с.
2. Доценко С.В. Случайные процессы в гидрофизических измерениях / С.В. Доценко. — Л.: Гидрометеоздат, 1983. — 239 с.

Поступила в редакцию 4.05.2007 г.

