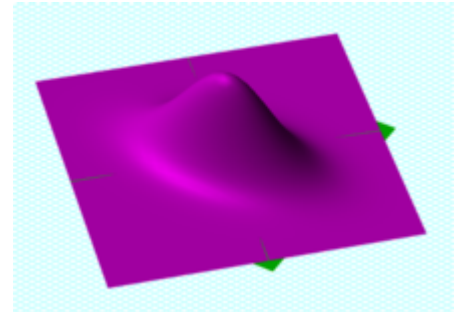


Covariance matrix

In probability theory and statistics, a **covariance matrix** (also known as **auto-covariance matrix**, **dispersion matrix**, **variance matrix**, or **variance–covariance matrix**) is a square matrix giving the covariance between each pair of elements of a given random vector. Any covariance matrix is symmetric and positive semi-definite and its main diagonal contains variances (i.e., the covariance of each element with itself).

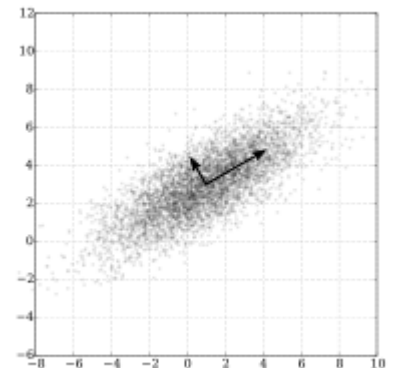
Intuitively, the covariance matrix generalizes the notion of variance to multiple dimensions. As an example, the variation in a collection of random points in two-dimensional space cannot be characterized fully by a single number, nor would the variances in the *x* and *y* directions contain all of the necessary information; a 2×2 matrix would be necessary to fully characterize the two-dimensional variation.

The covariance matrix of a random vector **X** is typically denoted by **K_{XX}** or **Σ**.



A bivariate Gaussian probability density function centered at (0, 0), with covariance matrix given by

$$\begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$



Sample points from a bivariate Gaussian distribution with a standard deviation of 3 in roughly the lower left–upper right direction and of 1 in the orthogonal direction. Because the *x* and *y* components co-vary, the variances of *x* and *y* do not fully describe the distribution. A 2×2 covariance matrix is needed; the directions of the arrows correspond to the eigenvectors of this covariance matrix and their lengths to the square roots of the eigenvalues.

Contents

Definition

[Conflicting nomenclatures and notations](#)

Properties

[Relation to the autocorrelation matrix](#)

[Relation to the correlation matrix](#)

[Inverse of the covariance matrix](#)

[Basic properties](#)

[Block matrices](#)

Partial covariance matrix

Covariance matrix as a parameter of a distribution

Covariance matrix as a linear operator

Which matrices are covariance matrices?

Complex random vectors

[Pseudo-covariance matrix](#)

Estimation

Applications

[Use in optimization](#)

[Covariance mapping](#)

[See also](#)[References](#)[Further reading](#)

Definition

Throughout this article, boldfaced unsubscripted \mathbf{X} and \mathbf{Y} are used to refer to random vectors, and unboldfaced subscripted X_i and Y_i are used to refer to scalar random variables.

If the entries in the column vector

$$\mathbf{X} = (X_1, X_2, \dots, X_n)^T$$

are random variables, each with finite variance and expected value, then the covariance matrix $\mathbf{K}_{\mathbf{X}\mathbf{X}}$ is the matrix whose (i, j) entry is the covariance^{[1]:p. 177}

$$\mathbf{K}_{X_i X_j} = \text{cov}[X_i, X_j] = \mathbf{E}[(X_i - \mathbf{E}[X_i])(X_j - \mathbf{E}[X_j])]$$

where the operator \mathbf{E} denotes the expected value (mean) of its argument.

Conflicting nomenclatures and notations

Nomenclatures differ. Some statisticians, following the probabilist William Feller in his two-volume book *An Introduction to Probability Theory and Its Applications*,^[2] call the matrix $\mathbf{K}_{\mathbf{X}\mathbf{X}}$ the **variance** of the random vector \mathbf{X} , because it is the natural generalization to higher dimensions of the 1-dimensional variance. Others call it the **covariance matrix**, because it is the matrix of covariances between the scalar components of the vector \mathbf{X} .

$$\text{var}(\mathbf{X}) = \text{cov}(\mathbf{X}, \mathbf{X}) = \mathbf{E}[(\mathbf{X} - \mathbf{E}[\mathbf{X}])(\mathbf{X} - \mathbf{E}[\mathbf{X}])^T].$$

Both forms are quite standard, and there is no ambiguity between them. The matrix $\mathbf{K}_{\mathbf{X}\mathbf{X}}$ is also often called the *variance-covariance matrix*, since the diagonal terms are in fact variances.

By comparison, the notation for the cross-covariance matrix *between* two vectors is

$$\text{cov}(\mathbf{X}, \mathbf{Y}) = \mathbf{K}_{\mathbf{X}\mathbf{Y}} = \mathbf{E}[(\mathbf{X} - \mathbf{E}[\mathbf{X}])(\mathbf{Y} - \mathbf{E}[\mathbf{Y}])^T].$$

Properties

Relation to the autocorrelation matrix

The auto-covariance matrix $\mathbf{K}_{\mathbf{X}\mathbf{X}}$ is related to the autocorrelation matrix $\mathbf{R}_{\mathbf{X}\mathbf{X}}$ by

$$\mathbf{K}_{\mathbf{X}\mathbf{X}} = \mathbf{E}[(\mathbf{X} - \mathbf{E}[\mathbf{X}])(\mathbf{X} - \mathbf{E}[\mathbf{X}])^T] = \mathbf{R}_{\mathbf{X}\mathbf{X}} - \mathbf{E}[\mathbf{X}] \mathbf{E}[\mathbf{X}]^T$$

where the autocorrelation matrix is defined as $\mathbf{R}_{\mathbf{X}\mathbf{X}} = \mathbf{E}[\mathbf{X}\mathbf{X}^T]$.

Relation to the correlation matrix

An entity closely related to the covariance matrix is the matrix of Pearson product-moment correlation coefficients between each of the random variables in the random vector \mathbf{X} , which can be written as

$$\text{corr}(\mathbf{X}) = (\text{diag}(\mathbf{K}_{\mathbf{X}\mathbf{X}}))^{-\frac{1}{2}} \mathbf{K}_{\mathbf{X}\mathbf{X}} (\text{diag}(\mathbf{K}_{\mathbf{X}\mathbf{X}}))^{-\frac{1}{2}},$$

where $\text{diag}(\mathbf{K}_{\mathbf{X}\mathbf{X}})$ is the matrix of the diagonal elements of $\mathbf{K}_{\mathbf{X}\mathbf{X}}$ (i.e., a diagonal matrix of the variances of X_i for $i = 1, \dots, n$).

Equivalently, the correlation matrix can be seen as the covariance matrix of the standardized random variables $X_i/\sigma(X_i)$ for $i = 1, \dots, n$.

$$\text{corr}(\mathbf{X}) = \begin{bmatrix} 1 & \frac{\mathbf{E}[(X_1-\mu_1)(X_2-\mu_2)]}{\sigma(X_1)\sigma(X_2)} & \dots & \frac{\mathbf{E}[(X_1-\mu_1)(X_n-\mu_n)]}{\sigma(X_1)\sigma(X_n)} \\ \frac{\mathbf{E}[(X_2-\mu_2)(X_1-\mu_1)]}{\sigma(X_2)\sigma(X_1)} & 1 & \dots & \frac{\mathbf{E}[(X_2-\mu_2)(X_n-\mu_n)]}{\sigma(X_2)\sigma(X_n)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\mathbf{E}[(X_n-\mu_n)(X_1-\mu_1)]}{\sigma(X_n)\sigma(X_1)} & \frac{\mathbf{E}[(X_n-\mu_n)(X_2-\mu_2)]}{\sigma(X_n)\sigma(X_2)} & \dots & 1 \end{bmatrix}.$$

Each element on the principal diagonal of a correlation matrix is the correlation of a random variable with itself, which always equals 1. Each off-diagonal element is between -1 and $+1$ inclusive.

Inverse of the covariance matrix

The inverse of this matrix, $\mathbf{K}_{\mathbf{X}\mathbf{X}}^{-1}$, if it exists, is the inverse covariance matrix, also known as the concentration matrix or precision matrix.^[3]

Basic properties

For $\mathbf{K}_{\mathbf{X}\mathbf{X}} = \text{var}(\mathbf{X}) = \mathbf{E}[(\mathbf{X} - \mathbf{E}[\mathbf{X}])(\mathbf{X} - \mathbf{E}[\mathbf{X}])^T]$ and $\mu_{\mathbf{X}} = \mathbf{E}[\mathbf{X}]$, where $\mathbf{X} = (X_1, \dots, X_n)^T$ is a n -dimensional random variable, the following basic properties apply:^[4]

1. $\mathbf{K}_{\mathbf{X}\mathbf{X}} = \mathbf{E}(\mathbf{X}\mathbf{X}^T) - \boldsymbol{\mu}_{\mathbf{X}}\boldsymbol{\mu}_{\mathbf{X}}^T$
2. $\mathbf{K}_{\mathbf{X}\mathbf{X}}$ is positive-semidefinite, i.e. $\mathbf{a}^T \mathbf{K}_{\mathbf{X}\mathbf{X}} \mathbf{a} \geq 0$ for all $\mathbf{a} \in \mathbb{R}^n$
3. $\mathbf{K}_{\mathbf{X}\mathbf{X}}$ is symmetric, i.e. $\mathbf{K}_{\mathbf{X}\mathbf{X}}^T = \mathbf{K}_{\mathbf{X}\mathbf{X}}$
4. For any constant (i.e. non-random) $m \times n$ matrix \mathbf{A} and constant $m \times 1$ vector \mathbf{a} , one has $\text{var}(\mathbf{A}\mathbf{X} + \mathbf{a}) = \mathbf{A} \text{var}(\mathbf{X}) \mathbf{A}^T$
5. If \mathbf{Y} is another random vector with the same dimension as \mathbf{X} , then $\text{var}(\mathbf{X} + \mathbf{Y}) = \text{var}(\mathbf{X}) + \text{cov}(\mathbf{X}, \mathbf{Y}) + \text{cov}(\mathbf{Y}, \mathbf{X}) + \text{var}(\mathbf{Y})$ where $\text{cov}(\mathbf{X}, \mathbf{Y})$ is the cross-covariance matrix of \mathbf{X} and \mathbf{Y} .

Block matrices

The joint mean $\boldsymbol{\mu}$ and joint covariance matrix $\boldsymbol{\Sigma}$ of \mathbf{X} and \mathbf{Y} can be written in block form

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_{\mathbf{X}} \\ \boldsymbol{\mu}_{\mathbf{Y}} \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \mathbf{K}_{\mathbf{X}\mathbf{X}} & \mathbf{K}_{\mathbf{X}\mathbf{Y}} \\ \mathbf{K}_{\mathbf{Y}\mathbf{X}} & \mathbf{K}_{\mathbf{Y}\mathbf{Y}} \end{bmatrix}$$

where $\mathbf{K}_{\mathbf{X}\mathbf{X}} = \text{var}(\mathbf{X})$, $\mathbf{K}_{\mathbf{Y}\mathbf{Y}} = \text{var}(\mathbf{Y})$ and $\mathbf{K}_{\mathbf{X}\mathbf{Y}} = \mathbf{K}_{\mathbf{Y}\mathbf{X}}^T = \text{cov}(\mathbf{X}, \mathbf{Y})$.

$\mathbf{K}_{\mathbf{X}\mathbf{X}}$ and $\mathbf{K}_{\mathbf{Y}\mathbf{Y}}$ can be identified as the variance matrices of the marginal distributions for \mathbf{X} and \mathbf{Y} respectively.

If \mathbf{X} and \mathbf{Y} are jointly normally distributed,

$$\mathbf{X}, \mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}),$$

then the conditional distribution for \mathbf{Y} given \mathbf{X} is given by

$$\mathbf{Y} | \mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{Y}|\mathbf{X}}, \mathbf{K}_{\mathbf{Y}|\mathbf{X}}),^{[5]}$$

defined by conditional mean

$$\boldsymbol{\mu}_{\mathbf{Y}|\mathbf{X}} = \boldsymbol{\mu}_{\mathbf{Y}} + \mathbf{K}_{\mathbf{Y}\mathbf{X}} \mathbf{K}_{\mathbf{X}\mathbf{X}}^{-1} (\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})$$

and conditional variance

$$\mathbf{K}_{\mathbf{Y}|\mathbf{X}} = \mathbf{K}_{\mathbf{Y}\mathbf{Y}} - \mathbf{K}_{\mathbf{Y}\mathbf{X}} \mathbf{K}_{\mathbf{X}\mathbf{X}}^{-1} \mathbf{K}_{\mathbf{X}\mathbf{Y}}.$$

The matrix $\mathbf{K}_{\mathbf{Y}\mathbf{X}} \mathbf{K}_{\mathbf{X}\mathbf{X}}^{-1}$ is known as the matrix of regression coefficients, while in linear algebra $\mathbf{K}_{\mathbf{Y}|\mathbf{X}}$ is the Schur complement of $\mathbf{K}_{\mathbf{X}\mathbf{X}}$ in $\boldsymbol{\Sigma}$.

The matrix of regression coefficients may often be given in transpose form, $\mathbf{K}_{\mathbf{X}\mathbf{X}}^{-1} \mathbf{K}_{\mathbf{X}\mathbf{Y}}$, suitable for post-multiplying a row vector of explanatory variables \mathbf{X}^T rather than pre-multiplying a column vector \mathbf{X} . In this form they correspond to the coefficients obtained by inverting the matrix of the normal equations of ordinary least squares (OLS).

Partial covariance matrix

A covariance matrix with all non-zero elements tells us that all the individual random variables are interrelated. This means that the variables are not only directly correlated, but also correlated via other variables indirectly. Often such indirect, common-mode correlations are trivial and uninteresting. They can be suppressed by calculating the partial covariance matrix, that is the part of covariance matrix that shows only the interesting part of correlations.

If two vectors of random variables \mathbf{X} and \mathbf{Y} are correlated via another vector \mathbf{I} , the latter correlations are suppressed in a matrix^[6]

$$\mathbf{K}_{\mathbf{XY}|\mathbf{I}} = \text{pcov}(\mathbf{X}, \mathbf{Y} | \mathbf{I}) = \text{cov}(\mathbf{X}, \mathbf{Y}) - \text{cov}(\mathbf{X}, \mathbf{I}) \text{cov}(\mathbf{I}, \mathbf{I})^{-1} \text{cov}(\mathbf{I}, \mathbf{Y}).$$

The partial covariance matrix $\mathbf{K}_{\mathbf{XY}|\mathbf{I}}$ is effectively the simple covariance matrix $\mathbf{K}_{\mathbf{XY}}$ as if the uninteresting random variables \mathbf{I} were held constant.

Covariance matrix as a parameter of a distribution

If a column vector \mathbf{X} of n possibly correlated random variables is jointly normally distributed, or more generally elliptically distributed, then its probability density function $f(\mathbf{X})$ can be expressed in terms of the covariance matrix Σ as follows^[6]

$$f(\mathbf{X}) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2} (\mathbf{X} - \mu)^T \Sigma^{-1} (\mathbf{X} - \mu)\right),$$

where $\mu = \mathbf{E}[\mathbf{X}]$ and $|\Sigma|$ is the determinant of Σ .

Covariance matrix as a linear operator

Applied to one vector, the covariance matrix maps a linear combination \mathbf{c} of the random variables \mathbf{X} onto a vector of covariances with those variables: $\mathbf{c}^T \Sigma = \text{cov}(\mathbf{c}^T \mathbf{X}, \mathbf{X})$. Treated as a bilinear form, it yields the covariance between the two linear combinations: $\mathbf{d}^T \Sigma \mathbf{c} = \text{cov}(\mathbf{d}^T \mathbf{X}, \mathbf{c}^T \mathbf{X})$. The variance of a linear combination is then $\mathbf{c}^T \Sigma \mathbf{c}$, its covariance with itself.

Similarly, the (pseudo-)inverse covariance matrix provides an inner product $\langle \mathbf{c} - \mu | \Sigma^+ | \mathbf{c} - \mu \rangle$, which induces the Mahalanobis distance, a measure of the "unlikelihood" of \mathbf{c} .

Which matrices are covariance matrices?

From the identity just above, let \mathbf{b} be a $(p \times 1)$ real-valued vector, then

$$\text{var}(\mathbf{b}^T \mathbf{X}) = \mathbf{b}^T \text{var}(\mathbf{X}) \mathbf{b},$$

which must always be nonnegative, since it is the variance of a real-valued random variable, so a covariance matrix is always a positive-semidefinite matrix.

The above argument can be expanded as follows:

$$\begin{aligned} \mathbf{w}^T \mathbf{E}[(\mathbf{X} - \mathbf{E}[\mathbf{X}])(\mathbf{X} - \mathbf{E}[\mathbf{X}])^T] \mathbf{w} &= \mathbf{E}[\mathbf{w}^T (\mathbf{X} - \mathbf{E}[\mathbf{X}])(\mathbf{X} - \mathbf{E}[\mathbf{X}])^T \mathbf{w}] \\ &= \mathbf{E}[(\mathbf{w}^T (\mathbf{X} - \mathbf{E}[\mathbf{X}]))^2] \geq 0, \end{aligned}$$

where the last inequality follows from the observation that $\mathbf{w}^T (\mathbf{X} - \mathbf{E}[\mathbf{X}])$ is a scalar.

Conversely, every symmetric positive semi-definite matrix is a covariance matrix. To see this, suppose \mathbf{M} is a $p \times p$ symmetric positive-semidefinite matrix. From the finite-dimensional case of the spectral theorem, it follows that \mathbf{M} has a nonnegative symmetric square root, which can be denoted by $\mathbf{M}^{1/2}$. Let \mathbf{X} be any $p \times 1$ column vector-valued random variable whose covariance matrix is the $p \times p$ identity matrix. Then

$$\text{var}(\mathbf{M}^{1/2} \mathbf{X}) = \mathbf{M}^{1/2} \text{var}(\mathbf{X}) \mathbf{M}^{1/2} = \mathbf{M}.$$

Complex random vectors

The variance of a complex scalar-valued random variable with expected value μ is conventionally defined using complex conjugation:

$$\text{var}(Z) = \mathbf{E}[(Z - \mu_Z) \overline{(Z - \mu_Z)}],$$

where the complex conjugate of a complex number z is denoted \bar{z} ; thus the variance of a complex random variable is a real number.

If $\mathbf{Z} = (Z_1, \dots, Z_n)^T$ is a column vector of complex-valued random variables, then the conjugate transpose \mathbf{Z}^H is formed by *both* transposing and conjugating. In the following expression, the product of a vector with its conjugate transpose results in a square matrix called the **covariance matrix**, as its expectation:^{[7]:p. 293}

$$\mathbf{K}_{\mathbf{Z}\mathbf{Z}} = \text{cov}[\mathbf{Z}, \mathbf{Z}] = \mathbf{E}[(\mathbf{Z} - \mu_{\mathbf{Z}})(\mathbf{Z} - \mu_{\mathbf{Z}})^H],$$

The matrix so obtained will be Hermitian positive-semidefinite,^[8] with real numbers in the main diagonal and complex numbers off-diagonal.

Properties

- The covariance matrix is a Hermitian matrix, i.e. $\mathbf{K}_{\mathbf{Z}\mathbf{Z}}^H = \mathbf{K}_{\mathbf{Z}\mathbf{Z}}$.^{[1]:p. 179}
- The diagonal elements of the covariance matrix are real.^{[1]:p. 179}

Pseudo-covariance matrix

For complex random vectors, another kind of second central moment, the **pseudo-covariance matrix** (also called **relation matrix**) is defined as follows:

$$\mathbf{J}_{\mathbf{Z}\mathbf{Z}} = \text{cov}[\mathbf{Z}, \overline{\mathbf{Z}}] = \mathbf{E}[(\mathbf{Z} - \mu_{\mathbf{Z}})(\mathbf{Z} - \mu_{\mathbf{Z}})^{\mathbf{T}}]$$

In contrast to the covariance matrix defined above, Hermitian transposition gets replaced by transposition in the definition. Its diagonal elements may be complex valued; it is a complex symmetric matrix.

Estimation

If $\mathbf{M}_{\mathbf{X}}$ and $\mathbf{M}_{\mathbf{Y}}$ are centred data matrices of dimension $p \times n$ and $q \times n$ respectively, i.e. with n columns of observations of p and q rows of variables, from which the row means have been subtracted, then, if the row means were estimated from the data, sample covariance matrices $\mathbf{Q}_{\mathbf{X}\mathbf{X}}$ and $\mathbf{Q}_{\mathbf{X}\mathbf{Y}}$ can be defined to be

$$\mathbf{Q}_{\mathbf{X}\mathbf{X}} = \frac{1}{n-1} \mathbf{M}_{\mathbf{X}} \mathbf{M}_{\mathbf{X}}^{\mathbf{T}}, \quad \mathbf{Q}_{\mathbf{X}\mathbf{Y}} = \frac{1}{n-1} \mathbf{M}_{\mathbf{X}} \mathbf{M}_{\mathbf{Y}}^{\mathbf{T}}$$

or, if the row means were known a priori,

$$\mathbf{Q}_{\mathbf{X}\mathbf{X}} = \frac{1}{n} \mathbf{M}_{\mathbf{X}} \mathbf{M}_{\mathbf{X}}^{\mathbf{T}}, \quad \mathbf{Q}_{\mathbf{X}\mathbf{Y}} = \frac{1}{n} \mathbf{M}_{\mathbf{X}} \mathbf{M}_{\mathbf{Y}}^{\mathbf{T}}.$$

These empirical sample covariance matrices are the most straightforward and most often used estimators for the covariance matrices, but other estimators also exist, including regularised or shrinkage estimators, which may have better properties.

Applications

The covariance matrix is a useful tool in many different areas. From it a transformation matrix can be derived, called a whitening transformation, that allows one to completely decorrelate the data or, from a different point of view, to find an optimal basis for representing the data in a compact way (see Rayleigh quotient for a formal proof and additional properties of covariance matrices). This is called principal component analysis (PCA) and the Karhunen–Loève transform (KL-transform).

The covariance matrix plays a key role in financial economics, especially in portfolio theory and its mutual fund separation theorem and in the capital asset pricing model. The matrix of covariances among various assets' returns is used to determine, under certain assumptions, the relative amounts of different assets that investors should (in a normative analysis) or are predicted to (in a positive analysis) choose to hold in a context of diversification.

Use in optimization

The evolution strategy, a particular family of Randomized Search Heuristics, fundamentally relies on a covariance matrix in its mechanism. The characteristic mutation operator draws the update step from a multivariate normal distribution using an evolving covariance matrix. There is a formal proof

that the evolution strategy's covariance matrix adapts to the inverse of the Hessian matrix of the search landscape, up to a scalar factor and small random fluctuations (proven for a single-parent strategy and a static model, as the population size increases, relying on the quadratic approximation).^[9] Intuitively, this result is supported by the rationale that the optimal covariance distribution can offer mutation steps whose equidensity probability contours match the level sets of the landscape, and so they maximize the progress rate.

Covariance mapping

In **covariance mapping** the values of the $\text{cov}(\mathbf{X}, \mathbf{Y})$ or $\text{pcov}(\mathbf{X}, \mathbf{Y} \mid \mathbf{I})$ matrix are plotted as a 2-dimensional map. When vectors \mathbf{X} and \mathbf{Y} are discrete random functions, the map shows statistical relations between different regions of the random functions. Statistically independent regions of the functions show up on the map as zero-level flatland, while positive or negative correlations show up, respectively, as hills or valleys.

In practice the column vectors \mathbf{X} , \mathbf{Y} , and \mathbf{I} are acquired experimentally as rows of n samples, e.g.

$$[\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n] = \begin{bmatrix} X_1(t_1) & X_2(t_1) & \cdots & X_n(t_1) \\ X_1(t_2) & X_2(t_2) & \cdots & X_n(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ X_1(t_m) & X_2(t_m) & \cdots & X_n(t_m) \end{bmatrix},$$

where $X_j(t_i)$ is the i -th discrete value in sample j of the random function $\mathbf{X}(t)$. The expected values needed in the covariance formula are estimated using the sample mean, e.g.

$$\langle \mathbf{X} \rangle = \frac{1}{n} \sum_{j=1}^n \mathbf{X}_j$$

and the covariance matrix is estimated by the sample covariance matrix

$$\text{cov}(\mathbf{X}, \mathbf{Y}) \approx \langle \mathbf{X}\mathbf{Y}^T \rangle - \langle \mathbf{X} \rangle \langle \mathbf{Y}^T \rangle,$$

where the angular brackets denote sample averaging as before except that the Bessel's correction should be made to avoid bias. Using this estimation the partial covariance matrix can be calculated as

$$\text{pcov}(\mathbf{X}, \mathbf{Y} \mid \mathbf{I}) = \text{cov}(\mathbf{X}, \mathbf{Y}) - \text{cov}(\mathbf{X}, \mathbf{I}) (\text{cov}(\mathbf{I}, \mathbf{I}) \backslash \text{cov}(\mathbf{I}, \mathbf{Y})),$$

where the backslash denotes the left matrix division operator, which bypasses the requirement to invert a matrix and is available in some computational packages such as Matlab.^[10]

Fig. 1 illustrates how a partial covariance map is constructed on an example of an experiment performed at the FLASH free-electron laser in

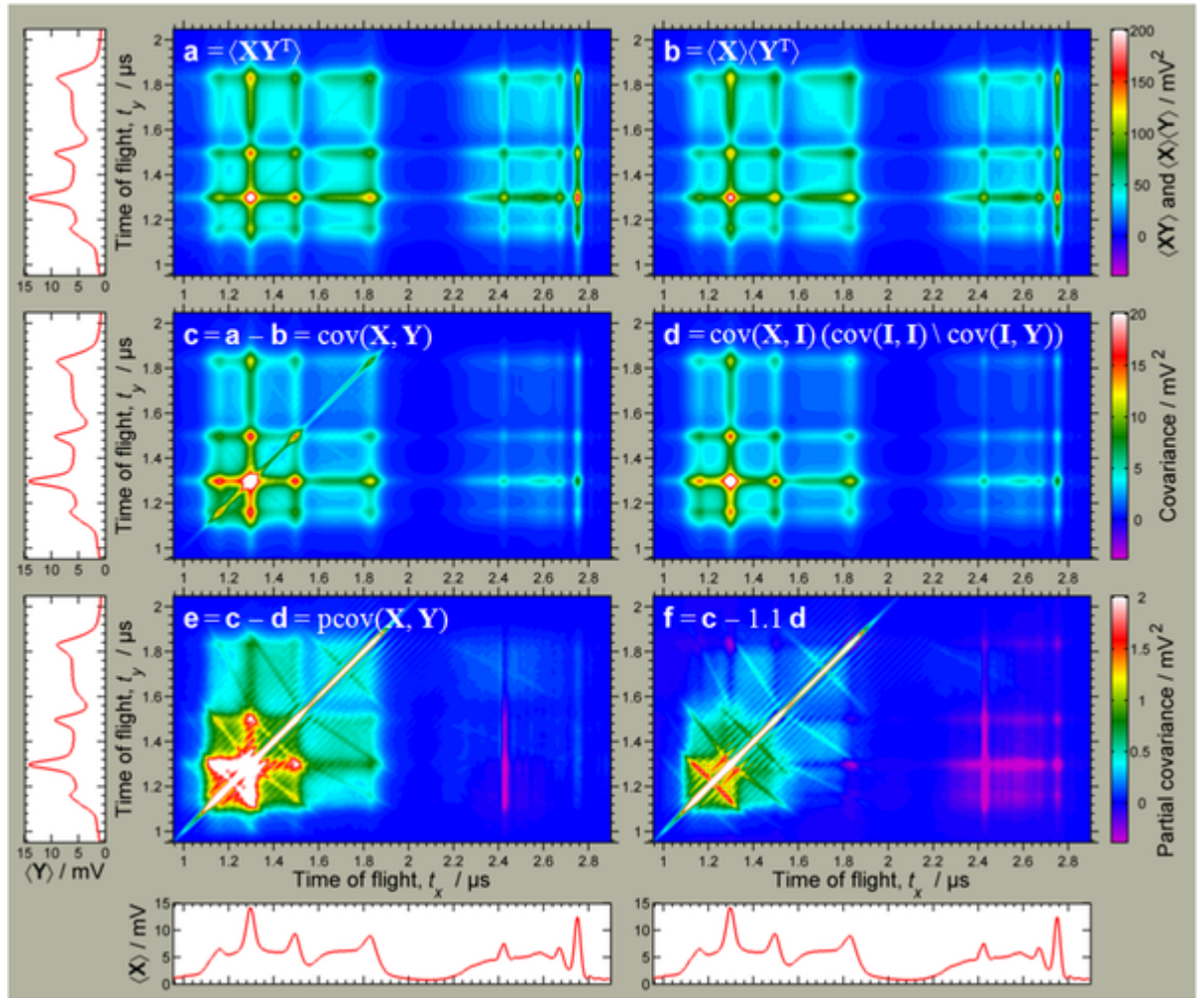


Figure 1: Construction of a partial covariance map of N_2 molecules undergoing Coulomb explosion induced by a free-electron laser.^[11] Panels **a** and **b** map the two terms of the covariance matrix, which is shown in panel **c**. Panel **d** maps common-mode correlations via intensity fluctuations of the laser. Panel **e** maps the partial covariance matrix that is corrected for the intensity fluctuations. Panel **f** shows that 10% overcorrection improves the map and makes ion-ion correlations clearly visible. Owing to momentum conservation these correlations appear as lines approximately perpendicular to the autocorrelation line (and to the periodic modulations which are caused by detector ringing).

Hamburg.^[11] The random function $\mathbf{X}(t)$ is the time-of-flight spectrum of ions from a Coulomb explosion of nitrogen molecules multiply ionised by a laser pulse. Since only a few hundreds of molecules are ionised at each laser pulse, the single-shot spectra are highly fluctuating. However, collecting typically $m = 10^4$ such spectra, $\mathbf{X}_j(t)$, and averaging them over j produces a smooth spectrum $\langle \mathbf{X}(t) \rangle$, which is shown in red at the bottom of Fig. 1. The average spectrum $\langle \mathbf{X} \rangle$ reveals several nitrogen ions in a form of peaks broadened by their kinetic energy, but to find the correlations between the ionisation stages and the ion momenta requires calculating a covariance map.

In the example of Fig. 1 spectra $\mathbf{X}_j(t)$ and $\mathbf{Y}_j(t)$ are the same, except that the range of the time-of-flight t differs. Panel **a** shows $\langle \mathbf{X}\mathbf{Y}^T \rangle$, panel **b** shows $\langle \mathbf{X} \rangle \langle \mathbf{Y}^T \rangle$ and panel **c** shows their difference, which is $\text{cov}(\mathbf{X}, \mathbf{Y})$ (note a change in the colour scale). Unfortunately, this map is overwhelmed by uninteresting, common-mode correlations induced by laser intensity fluctuating from shot to shot. To suppress such correlations the laser intensity I_j is recorded at every shot, put into \mathbf{I} and

$\text{pcov}(\mathbf{X}, \mathbf{Y} \mid \mathbf{I})$ is calculated as panels **d** and **e** show. The suppression of the uninteresting correlations is, however, imperfect because there are other sources of common-mode fluctuations than the laser intensity and in principle all these sources should be monitored in vector **I**. Yet in practice it is often sufficient to overcompensate the partial covariance correction as panel **f** shows, where interesting correlations of ion momenta are now clearly visible as straight lines centred on ionisation stages of atomic nitrogen.

Two-dimensional infrared spectroscopy

Two-dimensional infrared spectroscopy employs correlation analysis to obtain 2D spectra of the condensed phase. There are two versions of this analysis: synchronous and asynchronous. Mathematically, the former is expressed in terms of the sample covariance matrix and the technique is equivalent to covariance mapping.^[12]

See also

- Multivariate statistics
- Lewandowski-Kurowicka-Joe distribution
- Gramian matrix
- Eigenvalue decomposition
- Quadratic form (statistics)
- Principal components

References

1. Park, Kun Il (2018). *Fundamentals of Probability and Stochastic Processes with Applications to Communications*. Springer. ISBN 978-3-319-68074-3.
2. William Feller (1971). *An introduction to probability theory and its applications* (<https://books.google.com/books?id=K7kdAQAAMAAJ>). Wiley. ISBN 978-0-471-25709-7. Retrieved 10 August 2012.
3. Wasserman, Larry (2004). *All of Statistics: A Concise Course in Statistical Inference* (https://archive.org/details/springer_10.1007-978-0-387-21736-9). ISBN 0-387-40272-1.
4. Taboga, Marco (2010). "Lectures on probability theory and mathematical statistics" (<http://www.statlect.com/varian2.htm>).
5. Eaton, Morris L. (1983). *Multivariate Statistics: a Vector Space Approach*. John Wiley and Sons. pp. 116–117. ISBN 0-471-02776-6.
6. W J Krzanowski "Principles of Multivariate Analysis" (Oxford University Press, New York, 1988), Chap. 14.4; K V Mardia, J T Kent and J M Bibby "Multivariate Analysis (Academic Press, London, 1997), Chap. 6.5.3; T W Anderson "An Introduction to Multivariate Statistical Analysis" (Wiley, New York, 2003), 3rd ed., Chaps. 2.5.1 and 4.3.1.
7. Lapidoth, Amos (2009). *A Foundation in Digital Communication*. Cambridge University Press. ISBN 978-0-521-19395-5.
8. Brookes, Mike. "The Matrix Reference Manual" (<http://www.ee.ic.ac.uk/hp/staff/dmb/matrix/expect.html>).
9. Shir, O.M.; A. Yehudayoff (2020). "On the covariance-Hessian relation in evolution strategies" (<https://doi.org/10.1016/j.tcs.2019.09.002>). *Theoretical Computer Science*. Elsevier. **801**: 157–174. doi:10.1016/j.tcs.2019.09.002 (<https://doi.org/10.1016/j.tcs.2019.09.002>).

10. L J Frasiniski "Covariance mapping techniques" *J. Phys. B: At. Mol. Opt. Phys.* **49** 152004 (2016), [open access \(http://iopscience.iop.org/article/10.1088/0953-4075/49/15/152004\)](http://iopscience.iop.org/article/10.1088/0953-4075/49/15/152004)
11. O Kornilov, M Eckstein, M Rosenblatt, C P Schulz, K Motomura, A Rouzée, J Klei, L Foucar, M Siano, A Lübcke, F. Schapper, P Johnsson, D M P Holland, T Schlatholter, T Marchenko, S Düsterer, K Ueda, M J J Vrakking and L J Frasiniski "Coulomb explosion of diatomic molecules in intense XUV fields mapped by partial covariance" *J. Phys. B: At. Mol. Opt. Phys.* **46** 164028 (2013), [open access \(http://hdl.handle.net/10044/1/12267\)](http://hdl.handle.net/10044/1/12267)
12. I Noda "Generalized two-dimensional correlation method applicable to infrared, Raman, and other types of spectroscopy" *Appl. Spectrosc.* **47** 1329–36 (1993)

Further reading

- "Covariance matrix" (https://www.encyclopediaofmath.org/index.php?title=Covariance_matrix), *Encyclopedia of Mathematics*, EMS Press, 2001 [1994]
 - Weisstein, Eric W. "Covariance Matrix" (<https://mathworld.wolfram.com/CovarianceMatrix.html>). *MathWorld*.
 - van Kampen, N. G. (1981). *Stochastic processes in physics and chemistry* (<https://archive.org/details/stochasticproces0000kamp>). New York: North-Holland. ISBN 0-444-86200-5.
-

Retrieved from "https://en.wikipedia.org/w/index.php?title=Covariance_matrix&oldid=1095639672"

This page was last edited on 29 June 2022, at 13:55 (UTC).

Text is available under the Creative Commons Attribution-ShareAlike License 3.0; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.