协方差矩阵

wikipedia

15 October 2021

在概率论和统计学中，**协方差矩阵**（covariance matrix），也称为**自动协方差矩阵**（auto-covariance matrix）、**色散矩阵**（dispersion matrix）、**方差矩阵**（variance matrix）或**方差-协方差矩阵**（variance–covariance matrix），是一个平方矩阵，给出给定随机向量的每对元素之间的协方差。任意协方差矩阵都是对称的和半正定的矩阵，且其主对角线包含方差(即每个元素自身的协方差)。

直观地说，协方差矩阵将方差的概念推广到多维。例如，二维空间中随机点集合的变化不能完全用一个数字表示，在 x 和 y 方向上的变化也不能包含所有必要的信息；需要一个 2×2 矩阵来充分描述二维变化。

随机向量 $\mathbf{X}$ 的协方差矩阵通常由 $\mathrm{K}_{\mathbf{XX}}$ 或 $\mathbf{\Sigma}$ 表示。

1 定义

通过本文，粗体无下标的 $\mathbf{X}$ 和 $\mathbf{Y}$ 用于引用随机向量，非粗体有下标 X_i 和 Y_i 用于引用标量随机变量。

如果列向量的条目

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

是随机向量，每个变量都有有限的方差和期望值，那么协方差矩阵 $\mathrm{K}_{\mathbf{XX}}$ 是其 (i, j) 项为协方差的矩阵 [1]:p. 177

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

其中算子 E 表示其参数的期望值 (均值)。

换句话说，

$$\mathrm{K}_{\mathbf{XX}} = \begin{bmatrix} \mathrm{E}[(X_1 - \mathrm{E}[X_1])(X_1 - \mathrm{E}[X_1])] & \mathrm{E}[(X_1 - \mathrm{E}[X_1])(X_2 - \mathrm{E}[X_2])] & \cdots & \mathrm{E}[(X_1 - \mathrm{E}[X_1])(X_n - \mathrm{E}[X_n])] \\ \mathrm{E}[(X_2 - \mathrm{E}[X_2])(X_1 - \mathrm{E}[X_1])] & \mathrm{E}[(X_2 - \mathrm{E}[X_2])(X_2 - \mathrm{E}[X_2])] & \cdots & \mathrm{E}[(X_2 - \mathrm{E}[X_2])(X_n - \mathrm{E}[X_n])] \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{E}[(X_n - \mathrm{E}[X_n])(X_1 - \mathrm{E}[X_1])] & \mathrm{E}[(X_n - \mathrm{E}[X_n])(X_2 - \mathrm{E}[X_2])] & \cdots & \mathrm{E}[(X_n - \mathrm{E}[X_n])(X_n - \mathrm{E}[X_n])] \end{bmatrix}$$

上述定义等同于矩阵等式

$$\boxed{\mathrm{K}_{\mathbf{XX}} = \mathrm{cov}[\mathbf{X}, \mathbf{X}] = \mathrm{E}[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^{\mathrm{T}}] = \mathrm{E}[\mathbf{X}\mathbf{X}^{\mathrm{T}}] - \boldsymbol{\mu}_{\mathbf{X}}\boldsymbol{\mu}_{\mathbf{X}}^{\mathrm{T}}} \tag{1}$$

其中 $\boldsymbol{\mu}_{\mathbf{X}} = \mathrm{E}[\mathbf{X}]$。

1.1 方差的推广

这种形式，方程式 (1)，可视为标量值方差向更高维度的推广。记住，对于标量值随机变量 X

$$\sigma_X^2 = \mathrm{var}(X) = \mathrm{E}[(X - \mathrm{E}[X])^2] = \mathrm{E}[(X - \mathrm{E}[X]) \cdot (X - \mathrm{E}[X])].$$

实际上，自动协方差矩阵 $\mathrm{K}_{\mathbf{XX}}$ 的对角线上的条目是向量 $\mathbf{X}$ 的每个元素的方差。

1.2 相互冲突的术语与符号

在资料中所用名称各不相同。一些统计学家追随概率论学者 William Feller 在他的两卷书《概率论及其应用概论》[2] 中的观点，将矩阵 $\mathrm{K}_{\mathbf{XX}}$ 称为随机向量 $\mathbf{X}$ 的**方差 (variance)**，因为它是一维方差向更高维度的自然推广。其他人称之为**协方差矩阵 (covariance matrix)**，因为它是向量 $\mathbf{X}$ 的标量分量之间的协方差矩阵。

$$\mathrm{var}(\mathbf{X}) = \mathrm{cov}(\mathbf{X}) = \mathrm{E}\left[(\mathbf{X} - \mathrm{E}[\mathbf{X}])(\mathbf{X} - \mathrm{E}[\mathbf{X}])^{\mathrm{T}}\right].$$

这两种形式都非常标准,它们之间没有歧义。矩阵 $\mathrm{K}_{\mathbf{XX}}$ 也经常被称为**方差-协方差矩阵** (*variance-covariance matrix*)，因为对角项实际上是方差。

通过比较，两个向量之间的互协方差矩阵 (cross-covariance matrix) 的记号为

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

2 性质

2.1 与自动相关矩阵的关系

自动协方差矩阵 $\mathrm{K}_{\mathbf{XX}}$ 与自动相关矩阵 (autocorrelation matrix) $\mathrm{R}_{\mathbf{XX}}$ 的关系为

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

其中，自相关矩阵定义为 $\mathrm{R}_{\mathbf{XX}} = \mathrm{E}[\mathbf{X}\mathbf{X}^{\mathrm{T}}]$。

2.2 与相关矩阵的关系

与协方差矩阵密切相关的实体是随机向量 $\mathbf{X}$ 中每个随机变量之间的皮尔逊积矩相关系数矩阵，其可写为

$$\mathrm{corr}(\mathbf{X}) = \left(\mathrm{diag}(\mathrm{K}_{\mathbf{XX}})\right)^{-\frac{1}{2}} \mathrm{K}_{\mathbf{XX}} \left(\mathrm{diag}(\mathrm{K}_{\mathbf{XX}})\right)^{-\frac{1}{2}},$$

其中 $\mathrm{diag}(\mathrm{K}_{\mathbf{XX}})$ 是 $\mathrm{K}_{\mathbf{XX}}$ 的对角元素的矩阵 (即，对于 $i = 1, \ldots, n$，X_i 的方差的对角矩阵)。

等效地，对于 $i = 1, \ldots, n$，相关矩阵可视为标准化随机变量 $X_i/\sigma(X_i)$ 的协方差矩阵。

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

相关矩阵主对角线上的每个元素都是随机变量与其自身的相关性，它始终等于 1。每个非对角元素位于位于 -1 和 $+1$ 之间。

2.3 协方差矩阵的逆矩阵

该矩阵的逆矩阵，$\mathrm{K}_{\mathbf{XX}}^{-1}$，如果它存在，为逆协方差矩阵，也称为浓度矩阵或精度矩阵 [3]。

2.4 基本性质

对于 $\mathrm{K}_{\mathbf{XX}} = \operatorname{var}(\mathbf{X}) = \mathrm{E}\left[(\mathbf{X} - \mathrm{E}[\mathbf{X}])(\mathbf{X} - \mathrm{E}[\mathbf{X}])^{\mathrm{T}}\right]$ 和 $\boldsymbol{\mu}_{\mathbf{X}} = \mathrm{E}[\mathbf{X}]$，其中 $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ 是一个 n 维随机变量，以下基本属性适用 [4]：

1. $\mathrm{K}_{\mathbf{XX}} = \mathrm{E}(\mathbf{X}\mathbf{X}^{\mathrm{T}}) - \boldsymbol{\mu}_{\mathbf{X}}\boldsymbol{\mu}_{\mathbf{X}}^{\mathrm{T}}$
2. $\mathrm{K}_{\mathbf{XX}}$ 是半正定的，即 $\mathbf{a}^{\mathrm{T}}\mathrm{K}_{\mathbf{XX}}\mathbf{a} \geq 0$ ，对于所有的 $\mathbf{a} \in \mathbb{R}^n$。
3. $\mathrm{K}_{\mathbf{XX}}$ 是对称的，即 $\mathrm{K}_{\mathbf{XX}}^{\mathrm{T}} = \mathrm{K}_{\mathbf{XX}}$。
4. 对于任意常数 (即非随机) $m \times n$ 矩阵 $\mathbf{A}$ 和常数 $m \times 1$ 向量 $\mathbf{a}$ ，有 $\operatorname{var}(\mathbf{AX}+\mathbf{a}) = \mathbf{A}\operatorname{var}(\mathbf{X})\,\mathbf{A}^{\mathrm{T}}$。
5. 如果 $\mathbf{Y}$ 是另一个与 $\mathbf{X}$ 维数相同的随机向量，则 $\operatorname{var}(\mathbf{X}+\mathbf{Y}) = \operatorname{var}(\mathbf{X}) + \operatorname{cov}(\mathbf{X},\mathbf{Y}) + \operatorname{cov}(\mathbf{Y},\mathbf{X}) + \operatorname{var}(\mathbf{Y})$，其中 $\operatorname{cov}(\mathbf{X},\mathbf{Y})$ 是 $\mathbf{X}$ 和 $\mathbf{Y}$ 的互协方差矩阵。

2.5 块矩阵

$\mathbf{X}$ 和 $\mathbf{Y}$ 的联合均值 μ 和联合协方差矩阵 $\boldsymbol{\Sigma}$ 可写为

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_{\mathbf{X}} \\ \boldsymbol{\mu}_{\mathbf{Y}} \end{bmatrix}, \qquad \boldsymbol{\Sigma} = \begin{bmatrix} \mathrm{K}_{\mathbf{XX}} & \mathrm{K}_{\mathbf{XY}} \\ \mathrm{K}_{\mathbf{YX}} & \mathrm{K}_{\mathbf{YY}} \end{bmatrix}$$

其中 $\mathrm{K}_{\mathbf{XX}} = \operatorname{var}(\mathbf{X})$，$\mathrm{K}_{\mathbf{YY}} = \operatorname{var}(\mathbf{Y})$，并且 $\mathrm{K}_{\mathbf{XY}} = \mathrm{K}_{\mathbf{YX}}^{\mathrm{T}} = \operatorname{cov}(\mathbf{X},\mathbf{Y})$。

$\mathrm{K}_{\mathbf{XX}}$ 和 $\mathrm{K}_{\mathbf{YY}}$ 可以分别识别为 $\mathbf{X}$ 和 $\mathbf{Y}$ 的边际分布的方差矩阵。

如果 $\mathbf{X}$ 和 $\mathbf{Y}$ 是联合正态分布，

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

则给定 $\mathbf{X}$ 对于 $\mathbf{Y}$ 的条件分布如下所示 [5]：

$$\mathbf{Y} \mid \mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{Y}|\mathbf{X}}, \mathrm{K}_{\mathbf{Y}|\mathbf{X}}),$$

其由条件均值定义为

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

并且条件方差为

$$\mathrm{K}_{\mathbf{Y}|\mathbf{X}} = \mathrm{K}_{\mathbf{YY}} - \mathrm{K}_{\mathbf{YX}}\mathrm{K}_{\mathbf{XX}}^{-1}\mathrm{K}_{\mathbf{XY}}.$$

矩阵 $\mathrm{K}_{\mathbf{YX}}\mathrm{K}_{\mathbf{XX}}^{-1}$ 被称为回归系数矩阵，而在线性代数中，$\mathrm{K}_{\mathbf{Y}|\mathbf{X}}$ 是在 $\boldsymbol{\Sigma}$ 中的 $\mathrm{K}_{\mathbf{XX}}$ 的舒尔补。

回归系数矩阵通常可以转置形式给出，$\mathrm{K}_{\mathbf{XX}}^{-1}\mathrm{K}_{\mathbf{XY}}$，适用于解释变量 $\mathbf{X}^{\mathrm{T}}$ 的行向量的右乘，而不是列向量 $\mathbf{X}$ 的左乘。在这种形式中，它们相当于通过最小二乘法 (ordinary least squares, OLS) 的正则方程的逆矩阵而获得的系数。

3 偏协方差矩阵

所有非零元素的协方差矩阵告诉我们，所有单个随机变量都是相互关联的。这意味着变量不仅直接相关，还通过其他变量间接相关。通常，这种间接的共模关联是琐碎和无趣的。它们可以通过计算部分协方差矩阵来抑制，部分协方差矩阵仅显示相关的有趣部分。

如果随机变量 $\mathbf{X}$ 和 $\mathbf{Y}$ 的两个向量通过另一个向量 $\mathbf{I}$ 关联，则后一个关联在矩阵中被抑制 [6]

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

偏协方差矩阵 $\mathrm{K}_{\mathbf{XY|I}}$ 实际上是简单的协方差矩阵 $\mathrm{K}_{\mathbf{XY}}$，就好像无趣的随机变量 $\mathbf{I}$ 保持不变一样。

4 作为分布参数的协方差矩阵

如果 n 个可能相关的随机变量的列向量 $\mathbf{X}$ 为联合正态分布，或更一般的椭圆分布，则其概率密度函数 $\mathrm{f}(\mathbf{X})$ 可以用协方差矩阵 $\mathbf{\Sigma}$ 表示如下 [6]

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

其中 $\mu = \mathrm{E}[\mathbf{X}]$，并且 $|\mathbf{\Sigma}|$ 是 $\mathbf{\Sigma}$ 的行列式。

5 作为线性算子的协方差矩阵

应用于一个向量，协方差矩阵将随机变量 $\mathbf{X}$ 的线性组合 $\mathbf{c}$ 映射到具有这些变量的协方差向量：$\mathbf{c}^{\mathrm{T}}\mathbf{\Sigma} = \mathrm{cov}(\mathbf{c}^{\mathrm{T}}\mathbf{X},\mathbf{X})$。作为双线性形式处理，它产生两个线性组合之间的协方差：$\mathbf{d}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{c} = \mathrm{cov}(\mathbf{d}^{\mathrm{T}}\mathbf{X},\mathbf{c}^{\mathrm{T}}\mathbf{X})$。则线性组合的方差是 $\mathbf{c}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{c}$，是它与自身的协方差。

类似地，(伪) 逆协方差矩阵提供内积 $\langle c-\mu|\Sigma^{+}|c-\mu\rangle$，它产生了 Mahalanobis 距离，这是 c 的“非似然”的度量。

6 哪些矩阵是协方差矩阵?

根据上面的恒等式，令 $\mathbf{b}$ 是一个 $(p\times 1)$ 实数值向量，则

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

其结果必须总是非负的，因为它是实值随机变量的方差，所以协方差矩阵总是半正定矩阵。

上述论点可以展开如下：

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

其中，从观察结果得到的最后一个不等式 $w^{\mathrm{T}}(\mathbf{X}-\mathrm{E}[\mathbf{X}])$ 是一个标量。

相反，每个对称半正定矩阵都是协方差矩阵。为此，假设 $\mathbf{M}$ 是一个 $p \times p$ 对称半正定矩阵。从谱定理的有限维情况来看，$\mathbf{M}$ 有一个非负对称平方根，可以用 $\mathbf{M}^{1/2}$ 表示。设 $\mathbf{X}$ 为任意 $p \times 1$ 列向量值随机变量，其协方差矩阵为 $p \times p$ 单位矩阵。则

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

7 复随机向量

复数**标量值**随机变量与期望值 μ 的方差通常使用复共轭定义：

$$\operatorname{var}(Z) = \operatorname{E}\left[(Z - \mu_Z)\overline{(Z - \mu_Z)}\right],$$

其中复数 z 的复共轭表示为 $\overline{z}$；因此，复随机变量的方差是实数。

如果 $\mathbf{Z} = (Z_1, \ldots, Z_n)^{\mathrm{T}}$ 是复值随机变量的列向量，则共轭转置 $\mathbf{Z}^{\mathrm{H}}$ 由转置和共轭两部分构成。在下面的表达式中，向量与其共轭转置的乘积产生一个称为**协方差矩阵**的平方矩阵，正如其期望值 [7]:p. 293：

$$\operatorname{K}_{\mathbf{ZZ}} = \operatorname{cov}[\mathbf{Z}, \mathbf{Z}] = \operatorname{E}\left[(\mathbf{Z} - \boldsymbol{\mu}_{\mathbf{Z}})(\mathbf{Z} - \boldsymbol{\mu}_{\mathbf{Z}})^{\mathrm{H}}\right].$$

这样得到的矩阵将是厄米特半正定的[8]，实数在主对角线上，复数在非对角线上。

7.1 性质

- 协方差矩阵是厄米特矩阵，即 $\operatorname{K}_{\mathbf{ZZ}}^{\mathrm{H}} = \operatorname{K}_{\mathbf{ZZ}}$。[1]:p. 179
- 协方差矩阵的对角元素是实数。[1]:p. 179

7.2 伪协方差矩阵

对于复随机向量，另一种第二中心矩，**伪协方差矩阵** (**pseudo-covariance matrix**)，也称为**关系矩阵** (**relation matrix**)，定义如下：

$$\operatorname{J}_{\mathbf{ZZ}} = \operatorname{cov}[\mathbf{Z}, \overline{\mathbf{Z}}] = \operatorname{E}\left[(\mathbf{Z} - \boldsymbol{\mu}_{\mathbf{Z}})(\mathbf{Z} - \boldsymbol{\mu}_{\mathbf{Z}})^{\mathrm{T}}\right]$$

与上面定义的协方差矩阵不同，定义中的厄米转置被转置所取代。其对角线元素可以是复数；它是一个复对称矩阵。

8 估计

如果 $\mathbf{M}_{\mathbf{X}}$ 和 $\mathbf{M}_{\mathbf{Y}}$ 分别是维数为 $p \times n$ 和 $q \times n$ 的中心数据矩阵，也就是说，对于 p 行和 q 行变量的 n 列观测值，行均值已从中减去，那么，如果行均值是从数据中估计的，则样本协方差矩阵 $\mathbf{Q}_{\mathbf{XX}}$ 和 $\mathbf{Q}_{\mathbf{XY}}$ 可以定义为

$$\mathbf{Q}_{\mathbf{XX}} = \frac{1}{n-1}\mathbf{M}_{\mathbf{X}}\mathbf{M}_{\mathbf{X}}^{\mathrm{T}}, \qquad \mathbf{Q}_{\mathbf{XY}} = \frac{1}{n-1}\mathbf{M}_{\mathbf{X}}\mathbf{M}_{\mathbf{Y}}^{\mathrm{T}}$$

或者，如果行均值是先验的，

$$\mathbf{Q_{XX}} = \frac{1}{n}\mathbf{M_X}\mathbf{M_X}^{\mathrm{T}}, \qquad \mathbf{Q_{XY}} = \frac{1}{n}\mathbf{M_X}\mathbf{M_Y}^{\mathrm{T}}.$$

这些经验样本协方差矩阵是协方差矩阵最直接和最常用的估计量，但也存在其他估计量，包括正则或收缩估计量，它们可能具有更好的性质。

9 应用

协方差矩阵是许多不同领域的有用工具。从中可以导出一个变换矩阵，称为白化变换，它允许从一个完全消除数据的相关性的视角，或者从其它不同的视角，找到以紧凑方式表示数据的最佳基 (有关协方差矩阵的形式证明和附加属性，请参见瑞利商 (Rayleigh Quotient))。这被称为主成分分析 (principal component analysis, PCA) 和 Karhunen–Loève 变换 (Karhunen–Loève transform, KL-transform)。

协方差矩阵在金融经济学中起着关键作用，特别是在投资组合理论及其互助基金分离定理和资本资产定价模型中。在某些假设下，各种资产回报之间的协方差矩阵用于确定投资者应该 (在规范分析中) 或预计 (在实证分析中) 在多元化背景下选择持有的不同资产的相对金额。

9.1 协方差映射

在协方差映射中，$\mathrm{cov}(\mathbf{X}, \mathbf{Y})$ 或 $\mathrm{pcov}(\mathbf{X}, \mathbf{Y} \mid \mathbf{I})$ 矩阵绘制为二维图。当向量 $\mathbf{X}$ 和 $\mathbf{Y}$ 是离散随机函数时，该图显示随机函数不同区域之间的统计关系。统计上独立的函数区域在地图上显示为零级平地，而正相关或负相关分别显示为丘陵或山谷。

在实践中，列向量 $\mathbf{X}$、$\mathbf{Y}$ 和 $\mathbf{I}$ 是作为 n 个样本的行通过实验获得的，例如：

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

其中，$X_j(t_i)$ 是随机函数 $X(t)$ 的样本 j 中的第 i 个离散值。协方差公式中所需的期望值使用样本均值进行估计，例如：

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

并且利用样本协方差矩阵估计协方差矩阵

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

其中，角括号表示与之前一样的样本平均值，但应进行贝塞尔校正以避免偏差。使用该估计，可计算部分协方差矩阵，如下所示：

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

其中，反斜杠表示左矩阵除法算子，它绕过了矩阵求逆的要求，在一些计算包 (如Matlab) 中可用 [9]。

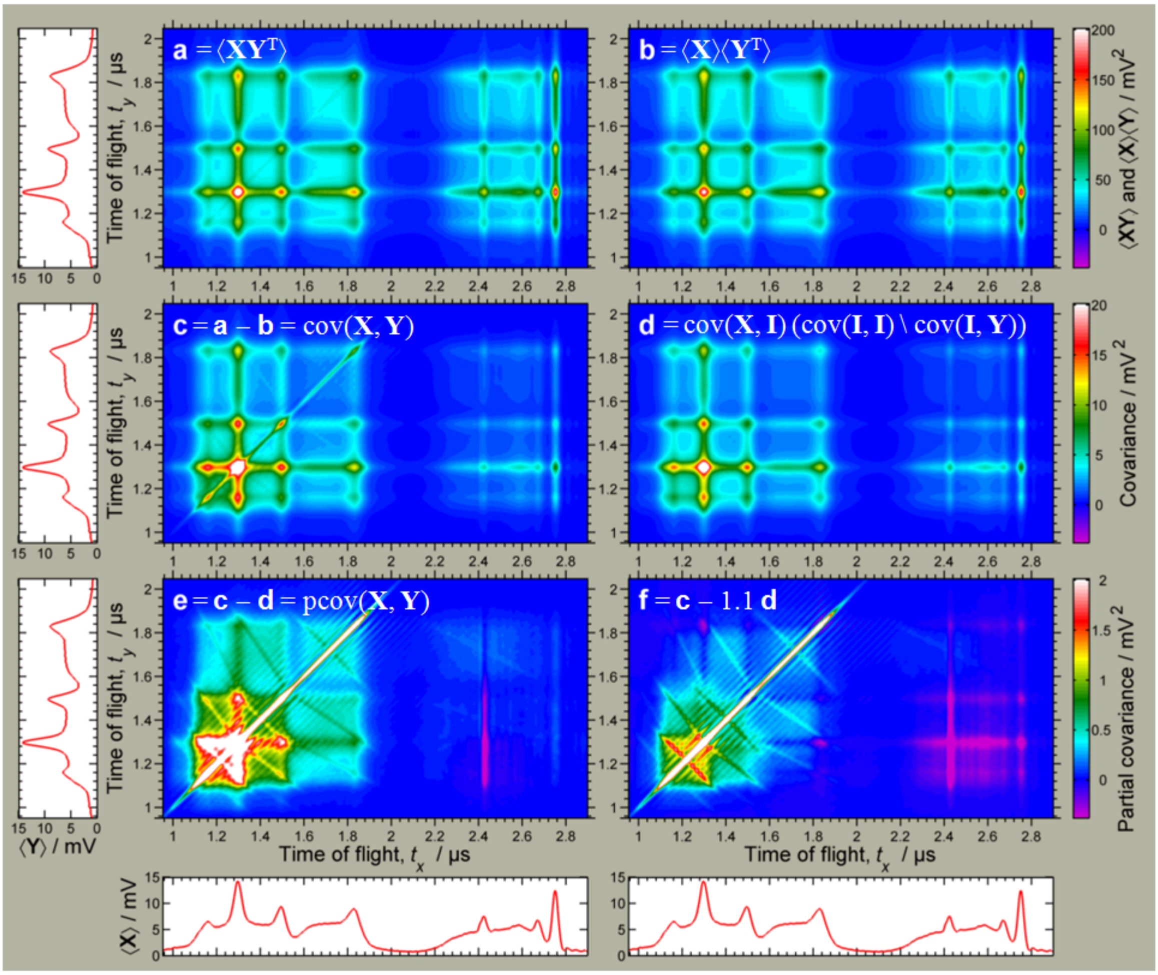

图 1: 构造自由电子激光诱导的经历库仑爆炸的 N_2 分子的部分协方差图 [10]。面板 **a** 和 **b** 绘制协方差矩阵的两项，如图 **c** 所示。面板 **d** 通过激光的强度波动绘制共模关联图。面板 **e** 映射针对强度波动校正的部分协方差矩阵。面板 **f** 显示 10% 的过度校正改善了图像，并使离子关联清晰可见。由于动量守恒，这些相关性显示为近似垂直于自相关线 (以及探测器振铃引起的周期调制) 的线。

图 1 说明了如何在 Hamburg 的 FLASH 自由电子激光器上进行的实验示例上构造部分协方差图 [10]。随机函数 $X(t)$ 是氮分子库仑爆炸产生的离子的飞行时间谱，被激光脉冲电离。由于每个激光脉冲只有几百个分子电离，因此单次激发的光谱高度波动。然而，典型地收集 $m = 10^4$ 这样的光谱，$\mathbf{X}_j(t)$，并在 j 上平均它们产生平滑的光谱 $\langle \mathbf{X}(t) \rangle$，在图 1 底部以红色显示。平均光谱 $\langle \mathbf{X} \rangle$ 以其动能加宽的峰形式揭示了几个氮离子，但要找到电离阶段和离子动量之间的相关性，需要计算协方差图。

在图 1 的示例中，光谱 $\mathbf{X}_j(t)$ 和 $\mathbf{Y}_j(t)$ 是相同的，只是飞行时间 t 的范围不同。面板 **a** 显示 $\langle \mathbf{X}\mathbf{Y}^{\mathrm{T}} \rangle$，面板 **b** 显示 $\langle \mathbf{X} \rangle \langle \mathbf{Y}^{\mathrm{T}} \rangle$，以及面板 **c** 显示了它们的差异，即 $\mathrm{cov}(\mathbf{X}, \mathbf{Y})$，注意色阶的变化。不幸的是，这张图被一个镜头到另一个镜头的激光强度波动引起的无趣的共模关联所淹没。为了抑制这种相关性，在每次拍摄时记录激光强度 I_j，并将其放入 $\mathbf{I}$ 中，并且 $\mathrm{pcov}(\mathbf{X}, \mathbf{Y} \mid \mathbf{I})$ 计算为面板 **d** 和 **e** 显示出来。然而，抑制无趣相关性是不完美的，因为除了激光强度之外还有其他共模波动源，原则上所有这些源都应该在向量 $\mathbf{I}$ 中进行监测。然而，在实践中，如面板 **f** 所示，过度补偿偏协方差校正通常是足够的，其中离子动量的有趣关联现在清晰可见，为以氮原子电离阶段为中心的直线。

9.2 二维红外光谱

二维红外光谱利用相关分析获得凝聚相的二维光谱。此分析有两种版本：同步和异步。在数学上，前者用样本协方差矩阵表示，该技术相当于协方差映射 [11]。

10 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. Wiley. ISBN 978-0-471-25709-7. Retrieved 10 August 2012.
3. Wasserman, Larry (2004). All of Statistics: A Concise Course in Statistical Inference. ISBN 0-387-40272-1.
4. Taboga, Marco (2010). "Lectures on probability theory and mathematical statistics".
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".
9. L J Frasinski "Covariance mapping techniques" J. Phys. B: At. Mol. Opt. Phys. 49 152004 (2016), open access
10. 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 Frasinski "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
11. I Noda "Generalized two-dimensional correlation method applicable to infrared, Raman, and other types of spectroscopy" Appl. Spectrosc. 47 1329–36 (1993)