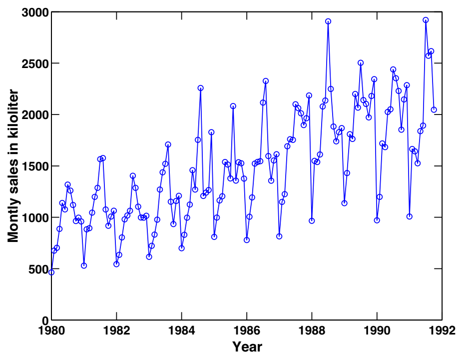
Stochastic vectors

Andreas Jakobsson

Stochastic vectors



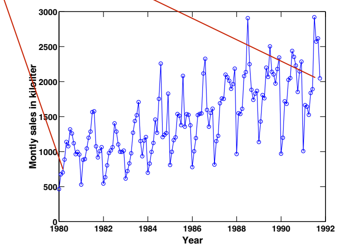
The monthly sales of red wine by Australian winemakers from 1980 to 1992.

Stochastic vectors

Let $\mathbf{x}$ denote a vector containing p stochastic variables, such that

$$\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_p \end{bmatrix}^T$$

where $(\cdot)^T$ and x_ℓ denote the transpose and the ℓth element of the vector $\mathbf{x}$, respectively.



Stochastic vectors

Let $\mathbf{x}$ denote a vector containing p stochastic variables, such that

$$\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_p \end{bmatrix}^T$$

where $(\cdot)^T$ and x_ℓ denote the transpose and the ℓth element of the vector $\mathbf{x}$, respectively.

This stochastic vector will follow the joint probability distribution function

$$F_{\mathbf{x}}(\alpha_1, \dots, \alpha_p) = P\{x_1 \leq \alpha_1, \dots, x_p \leq \alpha_p\}$$

where $P\{\cdot\}$ denotes the probability of the outcome $x_1 \leq \alpha_1, \dots, x_p \leq \alpha_p$.

Stochastic vectors

Let $\mathbf{x}$ denote a vector containing p stochastic variables, such that

$$\mathbf{x} = \begin{bmatrix} x_1 & \ldots & x_p \end{bmatrix}^T$$

where $(\cdot)^T$ and x_ℓ denote the transpose and the ℓth element of the vector $\mathbf{x}$, respectively.

This stochastic vector will follow the joint probability distribution function

$$F_{\mathbf{x}}(\alpha_1, \ldots, \alpha_p) = P\{x_1 \leq \alpha_1, \ldots, x_p \leq \alpha_p\}$$

where $P\{\cdot\}$ denotes the probability of the outcome $x_1 \leq \alpha_1, \ldots, x_p \leq \alpha_p$.

For a continuous sample space, the joint probability density function (PDF) is defined as

$$f_{\mathbf{x}}(\alpha_1, \ldots, \alpha_p) = \frac{\partial^p P\{x_1 \leq \alpha_1, \ldots, x_p \leq \alpha_p\}}{\partial \alpha_1, \ldots, \partial \alpha_p}$$

whereas for a discrete sample space, the joint PDF (or mass function) is

$$f_{\mathbf{x}}(\alpha_1, \ldots, \alpha_p) = P\{x_1 = \alpha_1, \ldots, x_p = \alpha_p\}$$

Stochastic vectors

A key concept for stochastic vectors is the *mean* value, defined as

$$\mathbf{m}_{\mathbf{x}} \equiv E\{\mathbf{x}\} = \begin{bmatrix} E\{x_1\} & \ldots & E\{x_p\} \end{bmatrix}^T$$

where $E\{\cdot\}$ denotes the statistical expectation, defined as

$$E\{g(\mathbf{x})\} = \int_{-\infty}^{\infty} g(\mathbf{x}) f(\mathbf{x}) \, d\mathbf{x}$$

where $g(\mathbf{x})$ denote some function of $\mathbf{x}$, and $f(\mathbf{x})$ the PDF of the vector.

Stochastic vectors

A key concept for stochastic vectors is the *mean* value, defined as

$$\mathbf{m}_{\mathbf{x}} \equiv E\{\mathbf{x}\} = \begin{bmatrix} E\{x_1\} & \ldots & E\{x_p\} \end{bmatrix}^T$$

where $E\{\cdot\}$ denotes the statistical expectation, defined as

$$E\{g(\mathbf{x})\} = \int_{-\infty}^{\infty} g(\mathbf{x}) f(\mathbf{x}) \, d\mathbf{x}$$

where $g(\mathbf{x})$ denote some function of $\mathbf{x}$, and $f(\mathbf{x})$ the PDF of the vector.

Furthermore, denote the *covariance matrix* of the vectors $\mathbf{x}$ and $\mathbf{y}$ by

$$\mathbf{R}_{\mathbf{x},\mathbf{y}} = C\{\mathbf{x}, \mathbf{y}\} = E\left\{[\mathbf{x} - \mathbf{m}_{\mathbf{x}}][\mathbf{y} - \mathbf{m}_{\mathbf{y}}]^*\right\} = E\left\{\mathbf{x}\mathbf{y}^*\right\} - \mathbf{m}_{\mathbf{x}}\mathbf{m}_{\mathbf{y}}^*$$

where $(\cdot)^*$ denotes the Hermitian, or conjugate transpose, and where the q-dimensional vectors $\mathbf{y}$ and $\mathbf{m}_{\mathbf{y}}$ are defined similarly to $\mathbf{x}$. Thus, $\mathbf{R}_{\mathbf{x},\mathbf{y}}$ is a $(p \times q)$-dimensional matrix with elements

$$\mathbf{R}_{\mathbf{x},\mathbf{y}} = \begin{bmatrix} C\{x_1, y_1\} & \ldots & C\{x_1, y_q\} \\ \vdots & \ddots & \vdots \\ C\{x_p, y_1\} & \ldots & C\{x_p, y_q\} \end{bmatrix}$$

Stochastic vectors

The (auto) covariance matrix $\mathbf{R}_{\mathbf{x},\mathbf{x}} \equiv \mathbf{R}_{\mathbf{x}}$ is *always*

(i) Hermitian, i.e., the matrix will satisfy $\mathbf{R}_{\mathbf{x}} = \mathbf{R}_{\mathbf{x}}^*$. If $\mathbf{x}$ is a real-valued vector, this implies that $\mathbf{R}_{\mathbf{x}} = \mathbf{R}_{\mathbf{x}}^T$; such matrices are termed *symmetric*.

(ii) positive semi-definite, here denoted $\mathbf{R}_{\mathbf{x}} \geq 0$, implying that $\mathbf{w}^*\mathbf{R}_{\mathbf{x}}\mathbf{w} \geq 0$, for all vectors $\mathbf{w}$. This also implies that the eigenvalues of $\mathbf{R}_{\mathbf{x}}$ are real-valued and non-negative.

Stochastic vectors

The (auto) covariance matrix $\mathbf{R}_{\mathbf{x},\mathbf{x}} \equiv \mathbf{R}_{\mathbf{x}}$ is *always*

(i) Hermitian, i.e., the matrix will satisfy $\mathbf{R}_{\mathbf{x}} = \mathbf{R}_{\mathbf{x}}^*$. If $\mathbf{x}$ is a real-valued vector, this implies that $\mathbf{R}_{\mathbf{x}} = \mathbf{R}_{\mathbf{x}}^T$; such matrices are termed *symmetric*.

(ii) positive semi-definite, here denoted $\mathbf{R}_{\mathbf{x}} \geq 0$, implying that $\mathbf{w}^*\mathbf{R}_{\mathbf{x}}\mathbf{w} \geq 0$, for all vectors $\mathbf{w}$. This also implies that the eigenvalues of $\mathbf{R}_{\mathbf{x}}$ are real-valued and non-negative.

Some further definitions:

(i) *independence*: The random variables in $\mathbf{x}$ are *independent*, if

$$f(\mathbf{x}) = \prod_{k=1}^{p} f(x_k)$$

(ii) *uncorrelated*: The random variables in $\mathbf{x}$ are *uncorrelated*, if

$$E\{\mathbf{x}\} = \prod_{k=1}^{p} E\{x_k\}$$

Conditional expectations

We are often interested in dependencies between different stochastic variables. One central notion to describe such dependencies can be described by the conditional distribution between the variables.

The *conditional density* of the random vector $\mathbf{y}$, given that $\mathbf{x} = \mathbf{x}_0$, for some value $\mathbf{x}_0$, is defined as

$$f_{\mathbf{y}|\mathbf{x}=\mathbf{x}_0}(\mathbf{y}) = \frac{f_{\mathbf{x},\mathbf{y}}(\mathbf{x}_0, \mathbf{y})}{f_{\mathbf{x}}(\mathbf{x}_0)} = \frac{f_{\mathbf{x},\mathbf{y}}(\mathbf{x}_0, \mathbf{y})}{\int f_{\mathbf{x},\mathbf{y}}(\mathbf{x}_0, \mathbf{y})\, d\mathbf{y}}$$

Conditional expectations

We are often interested in dependencies between different stochastic variables. One central notion to describe such dependencies can be described by the conditional distribution between the variables.

The *conditional density* of the random vector $\mathbf{y}$, given that $\mathbf{x} = \mathbf{x}_0$, for some value $\mathbf{x}_0$, is defined as

$$f_{\mathbf{y}|\mathbf{x}=\mathbf{x}_0}(\mathbf{y}) = \frac{f_{\mathbf{x},\mathbf{y}}(\mathbf{x}_0, \mathbf{y})}{f_{\mathbf{x}}(\mathbf{x}_0)} = \frac{f_{\mathbf{x},\mathbf{y}}(\mathbf{x}_0, \mathbf{y})}{\int f_{\mathbf{x},\mathbf{y}}(\mathbf{x}_0, \mathbf{y})\, d\mathbf{y}}$$

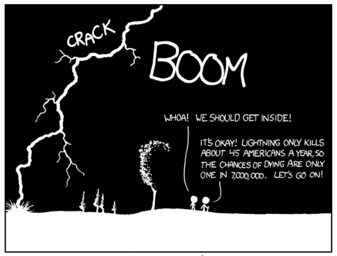
The conditional expectation is defined as

$$E\{\mathbf{y}|\mathbf{x} = \mathbf{x}_0\} = \int \mathbf{y} f_{\mathbf{y}|\mathbf{x}=\mathbf{x}_0}(\mathbf{y})\, d\mathbf{y}$$

To simplify notation, we use $E\{\mathbf{y}|\mathbf{x} = \mathbf{x}_0\} = E\{\mathbf{y}|\mathbf{x}\}$.

Note that if $\mathbf{x}$ and $\mathbf{y}$ are independent, then

$$E\{\mathbf{y}|\mathbf{x}\} = E\{\mathbf{y}\}$$



THE ANNUAL DEATH RATE AMONG PEOPLE WHO KNOW THAT STATISTIC IS ONE IN SIX.

Conditional expectations

Similarly, one can define the *conditional covariance matrix* as

$$C\{\mathbf{y}, \mathbf{z}|\mathbf{x}\} = E\left\{\left[\mathbf{y} - m_{\mathbf{y}|\mathbf{x}}\right]\left[\mathbf{z} - m_{\mathbf{z}|\mathbf{x}}\right]^* \Big| \mathbf{x}\right\}$$

where $m_{\mathbf{y}|\mathbf{x}} = E\{\mathbf{y}|\mathbf{x}\}$ and $m_{\mathbf{z}|\mathbf{x}} = E\{\mathbf{z}|\mathbf{x}\}$.

Conditional expectations

Similarly, one can define the *conditional covariance matrix* as

$$C\{\mathbf{y},\mathbf{z}|\mathbf{x}\} = E\left\{\left[\mathbf{y}-m_{\mathbf{y}|\mathbf{x}}\right]\left[\mathbf{z}-m_{\mathbf{z}|\mathbf{x}}\right]^{*}\Big|\mathbf{x}\right\}$$

where $m_{\mathbf{y}|\mathbf{x}} = E\{\mathbf{y}|\mathbf{x}\}$ and $m_{\mathbf{z}|\mathbf{x}} = E\{\mathbf{z}|\mathbf{x}\}$.

The *variance separation theorem* states that

$$V\{\mathbf{y}\} = E\Big\{V[\mathbf{y}|\mathbf{x}]\Big\} + V\Big\{E[\mathbf{y}|\mathbf{x}]\Big\}$$
$$C\{\mathbf{y},\mathbf{z}\} = E\Big\{C[\mathbf{y},\mathbf{z}|\mathbf{x}]\Big\} + C\Big\{E[\mathbf{y}|\mathbf{x}], E[\mathbf{z}|\mathbf{x}]\Big\}$$

where the expectations and (co)variances are taken with respect to the appropriate variables.

Conditional expectations

Similarly, one can define the *conditional covariance matrix* as

$$C\{\mathbf{y},\mathbf{z}|\mathbf{x}\} = E\left\{\left[\mathbf{y}-m_{\mathbf{y}|\mathbf{x}}\right]\left[\mathbf{z}-m_{\mathbf{z}|\mathbf{x}}\right]^{*}\Big|\mathbf{x}\right\}$$

where $m_{\mathbf{y}|\mathbf{x}} = E\{\mathbf{y}|\mathbf{x}\}$ and $m_{\mathbf{z}|\mathbf{x}} = E\{\mathbf{z}|\mathbf{x}\}$.

The *variance separation theorem* states that

$$V\{\mathbf{y}\} = E\Big\{V[\mathbf{y}|\mathbf{x}]\Big\} + V\Big\{E[\mathbf{y}|\mathbf{x}]\Big\}$$
$$C\{\mathbf{y},\mathbf{z}\} = E\Big\{C[\mathbf{y},\mathbf{z}|\mathbf{x}]\Big\} + C\Big\{E[\mathbf{y}|\mathbf{x}], E[\mathbf{z}|\mathbf{x}]\Big\}$$

where the expectations and (co)variances are taken with respect to the appropriate variables.

Example:

Consider $\mathbf{y} = a\mathbf{x} + \mathbf{e}$, where $\mathbf{x}$ and $\mathbf{e}$ are mutually independent, a is a real-valued constant, and the mean of $\mathbf{e}$ is zero. Then,

$$E\{\mathbf{y}|\mathbf{x}\} = E\{a\mathbf{x}+\mathbf{e}|\mathbf{x}\} = a\mathbf{x}$$
$$V\{\mathbf{y}|\mathbf{x}\} = V\{\mathbf{e}\}$$

Conditional expectations

Similarly, one can define the *conditional covariance matrix* as

$$C\{\mathbf{y},\mathbf{z}|\mathbf{x}\} = E\left\{\left[\mathbf{y}-m_{\mathbf{y}|\mathbf{x}}\right]\left[\mathbf{z}-m_{\mathbf{z}|\mathbf{x}}\right]^{*}\Big|\mathbf{x}\right\}$$

where $m_{\mathbf{y}|\mathbf{x}} = E\{\mathbf{y}|\mathbf{x}\}$ and $m_{\mathbf{z}|\mathbf{x}} = E\{\mathbf{z}|\mathbf{x}\}$.

The *variance separation theorem* states that

$$V\{\mathbf{y}\} = E\Big\{V[\mathbf{y}|\mathbf{x}]\Big\} + V\Big\{E[\mathbf{y}|\mathbf{x}]\Big\}$$
$$C\{\mathbf{y},\mathbf{z}\} = E\Big\{C[\mathbf{y},\mathbf{z}|\mathbf{x}]\Big\} + C\Big\{E[\mathbf{y}|\mathbf{x}], E[\mathbf{z}|\mathbf{x}]\Big\}$$

where the expectations and (co)variances are taken with respect to the appropriate variables.

Example:

Consider $\mathbf{y} = a\mathbf{x} + \mathbf{e}$, where $\mathbf{x}$ and $\mathbf{e}$ are mutually independent, a is a real-valued constant, and the mean of $\mathbf{e}$ is zero. Then,

$$E\{\mathbf{y}|\mathbf{x}\} = E\{a\mathbf{x}+\mathbf{e}|\mathbf{x}\} = a\mathbf{x}$$
$$V\{\mathbf{y}|\mathbf{x}\} = V\{\mathbf{e}\}$$

Similarly,

$$E\{\mathbf{y}\} = E\{E\{\mathbf{y}|\mathbf{x}\}\} = aE\{\mathbf{x}\}$$
$$V\{\mathbf{y}\} = E\Big\{V[\mathbf{y}|\mathbf{x}]\Big\} + V\Big\{E[\mathbf{y}|\mathbf{x}]\Big\} = V\{\mathbf{e}\} + a^2 V\{\mathbf{x}\}$$