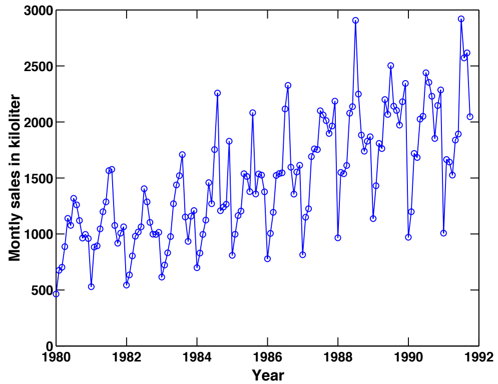
Time Series Analysis

Andreas Jakobsson

Contents:

- Multivariate random variables
- Conditional expectations
- Linear projections

Multivariate random variables



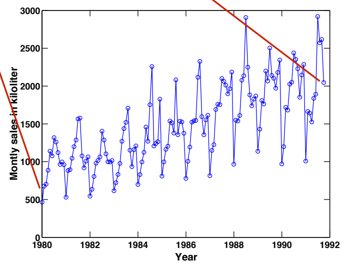
The monthly sales of red wine by Australian winemakers from 1980 to 1992.

Multivariate random variables

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.



Multivariate random variables

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

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

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

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

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

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

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

Multivariate random variables

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

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

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

$$\mathbf{R_{x,y}} = C\{\mathbf{x},\mathbf{y}\} = E\left\{[\mathbf{x} - \mathbf{m_x}]\,[\mathbf{y} - \mathbf{m_y}]^*\right\} = E\left\{\mathbf{x}\mathbf{y}^*\right\} - \mathbf{m_x}\mathbf{m_y}^*$$

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

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

Multivariate random variables

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

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

(ii) positive semi-definite, here denoted $\mathbf{R_x} \geq 0$, implying that $\mathbf{w}^*\mathbf{R_x}\mathbf{w} \geq 0$, for all vectors $\mathbf{w}$. This also implies that the eigenvalues of $\mathbf{R_x}$ 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

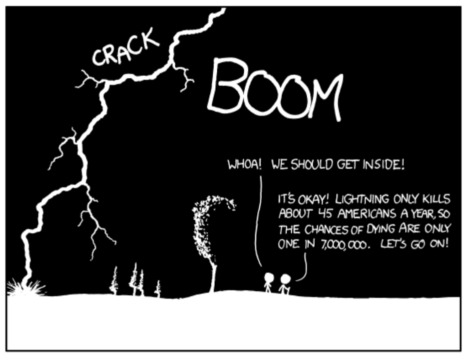
$$E\{\mathbf{x}\} = \prod_{k=1}^{p} E\{x_k\}$$

Multivariate random variables

If the $\mathbf{x}$ and $\mathbf{y}$ are independent, are they then also uncorrelated?

- Yes, independent variables are alway uncorrelated.
- No, there are independent variables that are correlated.

Conditional expectations



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

www.xkcd.com

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

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

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

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

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

Conditional expectations

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

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

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

The *variance separation theorem* states that

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

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

Conditional expectations

Is the following statement true?

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

- Yes, this is correct.
- No, this makes no sense.

Conditional expectations

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,

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

Similarly,

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

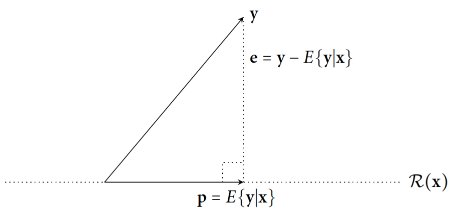
Linear projections

In this course, we will make use of conditional expectations to define the linear projection of one stochastic variable onto another.

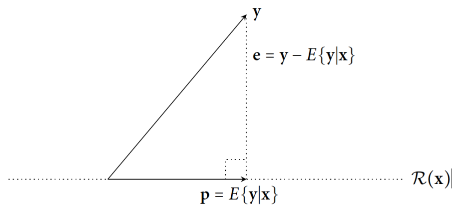
The *linear projection* of $\mathbf{y}$ onto the space spanned by $\mathbf{x}$, the so-called *range space*, denoted $\mathcal{R}(\mathbf{x})$, is defined as

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

where $\mathbf{a} \in \mathcal{R}(\mathbf{x})$ and $\mathbf{B}$ is a deterministic matrix of appropriate dimension.



Linear projections



The geometrical interpretation is quite helpful. For instance, from it, we can conclude the so-called *principle of orthogonality*, stating that

$$C\Big\{\mathbf{y} - E\{\mathbf{y}|\mathbf{x}\}, \mathbf{x}\Big\} = \mathbf{0}$$

That is, the error vector $\mathbf{e} = \mathbf{y} - E\{\mathbf{y}|\mathbf{x}\}$ is uncorrelated with $\mathbf{x}$.

Linear projections

Let $\mathbf{z}$ denote the concatenated vector

$$\mathbf{z} = \begin{bmatrix} \mathbf{x}^T & \mathbf{y}^T \end{bmatrix}^T$$

having mean $E\{\mathbf{z}\} = \begin{bmatrix} \mathbf{m}_\mathbf{x}^T & \mathbf{m}_\mathbf{y}^T \end{bmatrix}^T$ and covariance matrix

$$\mathbf{R}_\mathbf{z} = \begin{bmatrix} \mathbf{R}_\mathbf{x} & \mathbf{R}_{\mathbf{y},\mathbf{x}} \\ \mathbf{R}_{\mathbf{x},\mathbf{y}} & \mathbf{R}_\mathbf{y} \end{bmatrix}$$

Then, the linear projection of $\mathbf{y}$ onto $\mathbf{x}$, can be expressed as

$$E\{\mathbf{y}|\mathbf{x}\} = \mathbf{m}_\mathbf{y} + \mathbf{R}_{\mathbf{y},\mathbf{x}}\mathbf{R}_\mathbf{x}^{-1}(\mathbf{x} - \mathbf{m}_\mathbf{x})$$

This will be the optimal linear projection, i.e., the projection that yields the minimum prediction error variance among all linear projections. Furthermore, the difference $\mathbf{e} = \mathbf{y} - E\{\mathbf{y}|\mathbf{x}\}$ will have the variance

$$V\{\mathrm{e}|\mathbf{x}\} = \mathbf{R}_\mathbf{y} - \mathbf{R}_{\mathbf{y},\mathbf{x}}\mathbf{R}_\mathbf{x}^{-1}\mathbf{R}_{\mathbf{y},\mathbf{x}}^* = E\Big\{V\{\mathbf{y}|\mathbf{x}\}\Big\}$$

If $\mathbf{x}$ and $\mathbf{y}$ are Normal distributed, then $\mathbf{e}$ and $\mathbf{x}$ are independent; otherwise, they are uncorrelated.

Linear projections

The space orthogonal to the range space, $\mathcal{R}(\mathbf{A})$, is the

- The null space, $\mathcal{N}(\mathbf{A})$
- The left null space, $\mathcal{N}(\mathbf{A}^T)$
- The row space, $\mathcal{R}(\mathbf{A}^T)$

Linear projections

Let **a** denote the (monthly) measurement of the wine sales from 1980 to 1992, as shown earlier on, such that

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

with x_1 denoting the oldest sample, and x_p the most recent. As we wish to predict the *future* sales for the following month, x_{p+1}, we form

$$\mathbf{b} = \begin{bmatrix} x_2 & \dots & x_p & x_{p+1} \end{bmatrix}^T$$

Which of the following is then true?

- The projection of **a** onto **b** can be expressed without using x_{p+1}, and can, with some further information, be used to find the prediction of x_{p+1}.
- The projection of **b** onto **a** can be expressed without using x_{p+1}, and can, with some further information, be used to find the prediction of x_{p+1}.