Chapter 4
Jointly distributed Random variables

Continuous Multivariate distributions

Continuous Random Variables

Joint Probability Density Function (pdf)

Definition: Joint Probability density function

Two random variable are said to have *joint probability density function* $f(x, y)$ if

1. $f(x, y) \geq 0.$
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1.$
3. $P[\,(X, Y) \in A] = \int_A \int_A f(x, y)\, dx\, dy$

Marginal and Condition Density

Definition: Marginal Density

Let X and Y denote two RVs with joint pdf $f(x, y)$, then the *marginal density* of X is

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

and the *marginal density* of Y is $f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$

Definition: Conditional Density

Let X and Y denote two RVs with joint pdf $f(x, y)$ and marginal densities $f_X(x)$, $f_Y(y)$, then the *conditional density* of Y given $X = x$ and the conditional density of X given $Y = y$ are given by

$$f_{Y|X}\left(y\,|x\right) = \frac{f\left(x, y\right)}{f_X\left(x\right)} \qquad f_{X|Y}\left(x\,|y\right) = \frac{f\left(x, y\right)}{f_Y\left(y\right)}$$

Joint MGF

Definition: MGF of (X,Y)

Let X and Y be two RVs with joint pdf $f(x,y)$, then the MGF of X & Y:

$$m_{XY}(t_1,t_2) = E[\exp(t_1X + t_2Y)] = \iint_{R^2} \exp(t_1X + t_2Y)f(x,y)dxdy$$

Theorem:

The MGF of a pair of independent RVs is the product of the MGF of the corresponding marginal distributions. That is,

$$m_{XY}(t_1,t_2) = m_X(t_1)\, m_Y(t_2)$$

Proof:

$$m_{XY}(t_1,t_2) = \iint \exp(t_1X + t_2Y)f(x,y)dxdy$$

$$= \iint \exp(t_1X)\exp(t_2Y)f(x)f(y)dxdy$$

$$= \int \exp(t_1X)f(x)dx \int \exp(t_2Y)f(y)dy = m_X(t_1)m_Y(t_2)$$

Marginal MGF

Definition: MGF of the marginal distribution of X (and Y)

Let $m_{XY}(t_1,t_2)$ be the MGF of (X,Y), then the MGF of the marginal distributions of X and Y are, respectively, $m_{XY}(t_1,0)$ and $m_{XY}(0,t_2)$

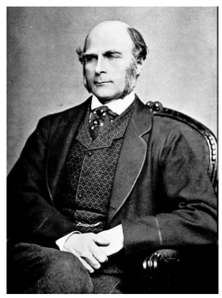
Proof:

$$m_X(t) = \int_{-\infty}^{\infty} \exp(tX)f_X(x)dx = \int_{-\infty}^{\infty} \exp(tX)[\int_{-\infty}^{\infty} f_{XY}(x,y)dy]dx =$$

$$= \iint \exp(tX)f_{XY}(x,y)dydx = m_{XY}(t,0)$$

Similar derivation for Y.

The bivariate Normal distribution



Sir Francis Galton (1822 –1911, England)

The bivariate normal distribution

Let the joint distribution be given by:

$$f\left(x_1, x_2\right) = \frac{1}{\left(2\pi\right)\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2}Q\left(x_1, x_2\right)}$$

where

$$Q\left(x_1, x_2\right) = \frac{\left\{\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right\}}{1-\rho^2}$$

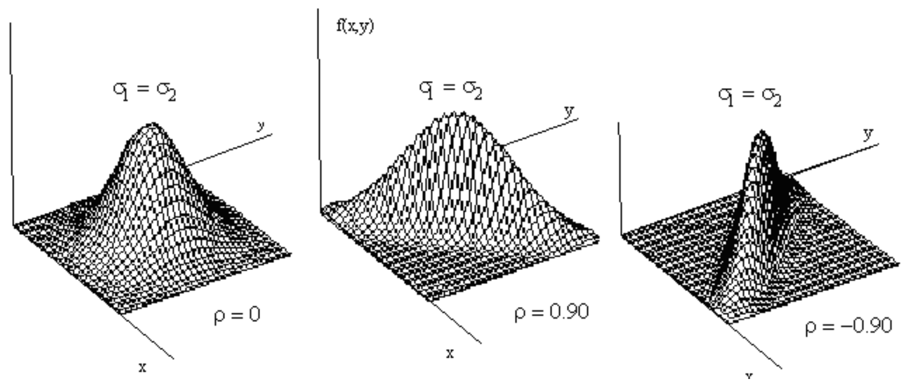
This distribution is called the *bivariate Normal distribution*.

The parameters are μ_1, μ_2, σ_1, σ_2 and ρ.

The properties of this distribution were studied by Francis Galton and discovered its relation to the regression, term Galton coined.

The bivariate normal distribution

Surface Plots of the bivariate Normal distribution



The bivariate normal distribution

<u>Note</u>: We can have a more compact joint using linear algebra:

$$f(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}}\exp\left[-\frac{1}{2(1-\rho^2)}\left((\frac{x_1-\mu_1}{\sigma_1})^2-2\rho(\frac{x_1-\mu_1}{\sigma_1})(\frac{x_2-\mu_2}{\sigma_2})+(\frac{x_2-\mu_2}{\sigma_2})^2\right)^2\right]$$

$$=\frac{1}{2\pi(|\Sigma|)^{1/2}}\exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

(1) Determine the inverse and determinant of Σ (the covariance matrix)

$$\Sigma=\begin{bmatrix}\sigma_1^2 & \sigma_{21}\\ \sigma_{12} & \sigma_2^2\end{bmatrix}\quad\Rightarrow|\Sigma|=\sigma_1^2\sigma_2^2-\sigma_{12}^2=\sigma_1^2\sigma_2^2(1-\frac{\sigma_{12}^2}{\sigma_1^2\sigma_2^2})=\sigma_1^2\sigma_2^2(1-\rho^2)$$

$$\Rightarrow\Sigma^{-1}=\frac{1}{|\Sigma|}\begin{bmatrix}\sigma_2^2 & -\sigma_{21}\\ -\sigma_{12} & \sigma_1^2\end{bmatrix}$$

The bivariate normal distribution

(2) Write a quadratic form for $Q(x_1,x_2)$:

$$Q\left(x_1, x_2\right) = \frac{\left\{\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right\}}{1-\rho^2}$$

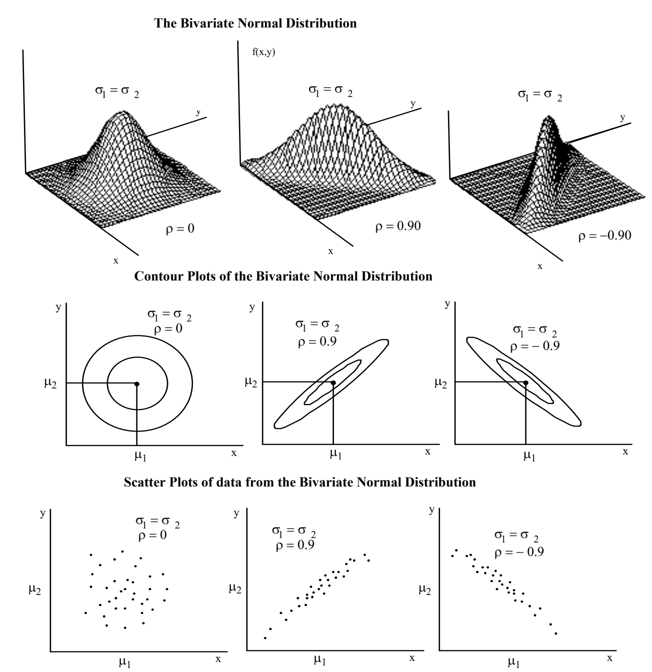
$$Q(x_1,x_2) = (\mathbf{x}-\boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})$$

$$= \left[(x_1-\mu_1) \quad (x_2-\mu_2)\right]\frac{1}{|\Sigma|}\begin{bmatrix} \sigma_2^2 & -\sigma_{21} \\ -\sigma_{12} & \sigma_1^2 \end{bmatrix}\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix} =$$

$$= \frac{1}{|\Sigma|}\left[(x_1-\mu_1)\sigma_2^2 - (x_2-\mu_2)\sigma_{12} \quad (x_1-\mu_1)\sigma_{21} - (x_2-\mu_2)\sigma_1^2\right]\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}$$

$$= \frac{1}{|\Sigma|}((x_1-\mu_1)^2\sigma_2^2 - 2\sigma_{21}(x_1-\mu_1)(x_2-\mu_2) + (x_2-\mu_2)^2\sigma_1^2)$$

$$= \frac{((x_1-\mu_1)^2\sigma_2^2 - 2\sigma_{21}(x_1-\mu_1)(x_2-\mu_2) + (x_2-\mu_2)^2\sigma_1^2)}{\sigma_1^2\sigma_2^2(1-\rho^2)}$$



The bivariate normal distribution: MGF

Let MGF of a bivariate normal is given by:

$$m_{XY}(t_1,t_2) = \exp[t_1\mu_X + t_2\mu_Y - \frac{1}{2}(t_1^2\sigma_X^2 + t_2^2\sigma_Y^2 + 2\rho_{XY}t_1t_2\sigma_X\sigma_Y)]$$

Note: When $\rho_{XY} = 0$ –i.e., X and Y are independent. The MGF is:

$$m_{XY}(t_1,t_2) = \exp[t_1\mu_X + t_2\mu_Y - \frac{1}{2}(t_1^2\sigma_X^2 + t_2^2\sigma_Y^2)]$$

Marginal distributions for the Bivariate Normal

Recall the definition of marginal distributions for continuous RV:

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1,x_2)\,dx_2 \quad \text{and} \quad f_2(x_2) = \int_{-\infty}^{\infty} f(x_1,x_2)\,dx_1$$

In the case of the bivariate normal distribution the marginal distribution of x_i is Normal with mean μ_i and standard deviation σ_i.

Proof:

The marginal distributions of x_2 is

$$f_2(x_2) = \int_{-\infty}^{\infty} f(x_1,x_2)\,dx_1 = \frac{1}{(2\pi)\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}Q(x_1,x_2)}\,dx_1$$

where

$$Q(x_1,x_2) = \frac{\left\{\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right\}}{1-\rho^2}$$

Marginal distributions for the Bivariate Normal

Now:

$$Q(x_1,x_2)=\frac{\left\{\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right\}}{1-\rho^2}$$

$$=\left(\frac{x_1-a}{b}\right)^2+c=\frac{x_1^2}{b^2}-2\frac{a}{b^2}x_1+\frac{a^2}{b^2}+c$$

$$=\frac{x_1^2}{\sigma_1^2\left(1-\rho^2\right)}-2\left[\frac{\mu_1}{\sigma_1^2\left(1-\rho^2\right)}+\rho\frac{x_2-\mu_2}{\sigma_2\sigma_1\left(1-\rho^2\right)}\right]x_1$$

$$+\quad\frac{\mu_1^2}{\sigma_1^2\left(1-\rho^2\right)}+2\rho\frac{\left(x_2-\mu_2\right)}{\sigma_2\sigma_1\left(1-\rho^2\right)}\mu_1+\frac{\left(x_2-\mu_2\right)^2}{\sigma_2^2\left(1-\rho^2\right)}$$

Marginal distributions for the Bivariate Normal

Hence $b^2=\sigma_1^2\left(1-\rho^2\right)$ or $b=\sigma_1\sqrt{1-\rho^2}$

Also

$$\frac{a}{b^2}=\frac{\mu_1}{\sigma_1^2\left(1-\rho^2\right)}+\rho\frac{x_2-\mu_2}{\sigma_2\sigma_1\left(1-\rho^2\right)}$$

$$=\frac{1}{\sigma_1^2\left(1-\rho^2\right)}\left[\mu_1+\rho\frac{\sigma_1}{\sigma_2}\left(x_2-\mu\right)\right]$$

and

$$a=\mu_1+\rho\frac{\sigma_1}{\sigma_2}\left(x_2-\mu\right)$$

Marginal distributions for the Bivariate Normal

Finally

$$\frac{a^2}{b^2} + c = \frac{\mu_1^2}{\sigma_1^2\left(1-\rho^2\right)} + 2\rho\frac{\left(x_2-\mu_2\right)}{\sigma_2\sigma_1\left(1-\rho^2\right)}\mu_1 + \frac{\left(x_2-\mu_2\right)^2}{\sigma_2^2\left(1-\rho^2\right)}$$

$$c = \frac{\mu_1^2}{\sigma_1^2\left(1-\rho^2\right)} + 2\rho\frac{\left(x_2-\mu_2\right)}{\sigma_2\sigma_1\left(1-\rho^2\right)}\mu_1 + \frac{\left(x_2-\mu_2\right)^2}{\sigma_2^2\left(1-\rho^2\right)} - \frac{a^2}{b^2}$$

$$= \frac{\mu_1^2}{\sigma_1^2\left(1-\rho^2\right)} + 2\rho\frac{\left(x_2-\mu_2\right)}{\sigma_2\sigma_1\left(1-\rho^2\right)}\mu_1 + \frac{\left(x_2-\mu_2\right)^2}{\sigma_2^2\left(1-\rho^2\right)} - \frac{\left[\mu_1 + \rho\frac{\sigma_1}{\sigma_2}\left(x_2-\mu\right)\right]^2}{\sigma_1^2\left(1-\rho^2\right)}$$

Marginal distributions for the Bivariate Normal

and

$$c = \frac{1}{\sigma_1^2\left(1-\rho^2\right)}\left[\mu_1^2 + 2\rho\frac{\sigma_1}{\sigma_2}\mu_1\left(x_2-\mu_2\right) + \frac{\sigma_1^2}{\sigma_2^2}\left(x_2-\mu_2\right)^2 - \left[\mu_1 + \rho\frac{\sigma_1}{\sigma_2}\left(x_2-\mu_2\right)\right]^2\right]$$

$$= \frac{1}{\sigma_1^2\left(1-\rho^2\right)}\left[\frac{\sigma_1^2}{\sigma_2^2}\left(1-\rho^2\right)\left(x_2-\mu_2\right)^2\right]$$

$$= \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2$$

Marginal distributions for the Bivariate Normal

Summarizing

$$Q(x_1,x_2)=\frac{\left\{\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right\}}{1-\rho^2}$$

$$=\left(\frac{x_1-a}{b}\right)^2+c$$

where $b=\sigma_1\sqrt{1-\rho^2}$

$$a=\mu_1+\rho\frac{\sigma_1}{\sigma_2}(x_2-\mu)$$

and $c=\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2$

Marginal distributions for the Bivariate Normal

Thus $f_2(x_2)=\int_{-\infty}^{\infty} f(x_1,x_2)\,dx_1$

$$=\frac{1}{(2\pi)\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{1}{2}Q(x_1,x_2)}\,dx_1$$

$$=\frac{1}{(2\pi)\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{1}{2}\left[\left(\frac{x_1-a}{b}\right)^2+c\right]}\,dx_1$$

$$=\frac{\sqrt{2\pi}\,b\,e^{-c/2}}{(2\pi)\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\,b}\mathrm{e}^{-\frac{1}{2}\left(\frac{x_1-a}{b}\right)^2}\,dx_1$$

$$=\frac{1}{\sqrt{2\pi}\,\sigma_2}e^{-\frac{1}{2}\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2}$$

Marginal distributions for the Bivariate Normal

• Thus the marginal distribution of x_2 is Normal with mean μ_2 and standard deviation σ_2.

• Similarly, the marginal distribution of x_1 is Normal with mean μ_1 and standard deviation σ_1.

Note: This derivation is much easier using MGFs.

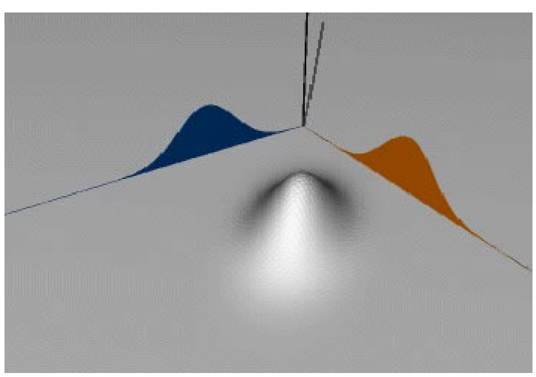
Use the MGF of a bivariate normal. To get the MGF of the marginal of X, set t_2=0.

$$m_{XY}(t_1,t_2) = \exp[t_1\mu_X + t_2\mu_Y - \frac{1}{2}(t_1^2\sigma_X^2 + t_2^2\sigma_Y^2 + 2\rho_{XY}t_1t_2\sigma_X\sigma_Y)]$$

$$m_{XY}(t_1,0) = \exp[t_1\mu_X - \frac{1}{2}(t_1^2\sigma_X^2)] = m_X(t_1)$$

Marginal distributions: Bivariate Normal

Bivariate Normal Distribution with marginal distributions



Conditional distributions for the Bivariate Normal

Recall the definition of conditional distributions for continuous RVs:

$$f_{1|2}\left(x_1 \middle| x_2\right) = \frac{f\left(x_1, x_2\right)}{f_2\left(x_2\right)} \quad \text{and} \quad f_{2|1}\left(x_2 \middle| x_1\right) = \frac{f\left(x_1, x_2\right)}{f_1\left(x_1\right)}$$

In the case of the bivariate normal distribution the conditional distribution of x_i given x_j is Normal with mean and standard deviation:

$$\mu_{i|j} = \mu_i + \rho \frac{\sigma_i}{\sigma_j}\left(x_j - \mu_j\right) \quad \text{and} \quad \sigma_{i|j} = \sigma_i \sqrt{1-\rho^2}$$

Conditional distributions: Bivariate Normal

Proof: |

$$f_{1|2} = \frac{f(x_1, x_2)}{f(x_2)} = \left. \frac{e^{-\frac{1}{2}Q(x_1,x_2)}}{(2\pi)\sigma_1\sigma_2\sqrt{1-\rho^2}} \middle/ \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2}\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2} \right.$$

$$= \frac{e^{-\frac{1}{2}Q(x_1,x_2)+\frac{1}{2}\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2}}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}} = \frac{e^{-\frac{1}{2}\left[\left(\frac{x_1-a}{b}\right)^2+c\right]+\frac{1}{2}\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2}}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}}$$

where

$$b = \sigma_1\sqrt{1-\rho^2} \qquad a = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2 - \mu_2) \qquad c = \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2$$

Conditional distributions: Bivariate Normal

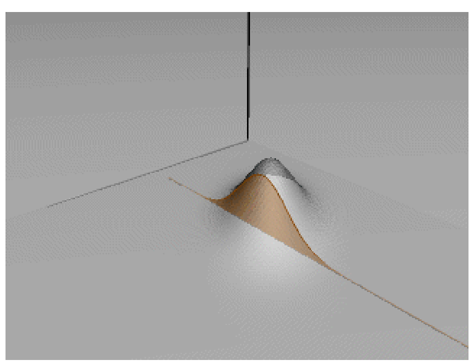
Hence $$f_{1|2}\left(x_1 \middle| x_2\right) = \frac{1}{\sqrt{2\pi}b}\, e^{-\frac{1}{2}\left(\frac{x_1-a}{b}\right)^2}$$

Then, the conditional distribution of x_2 given x_1 is Normal with mean and standard deviation:

$$a = \mu_{1|2} = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}\left(x_2 - \mu_2\right) \quad \text{and} \quad b = \sigma_{1|2} = \sigma_1\sqrt{1-\rho^2}$$

Conditional distributions: Bivariate Normal

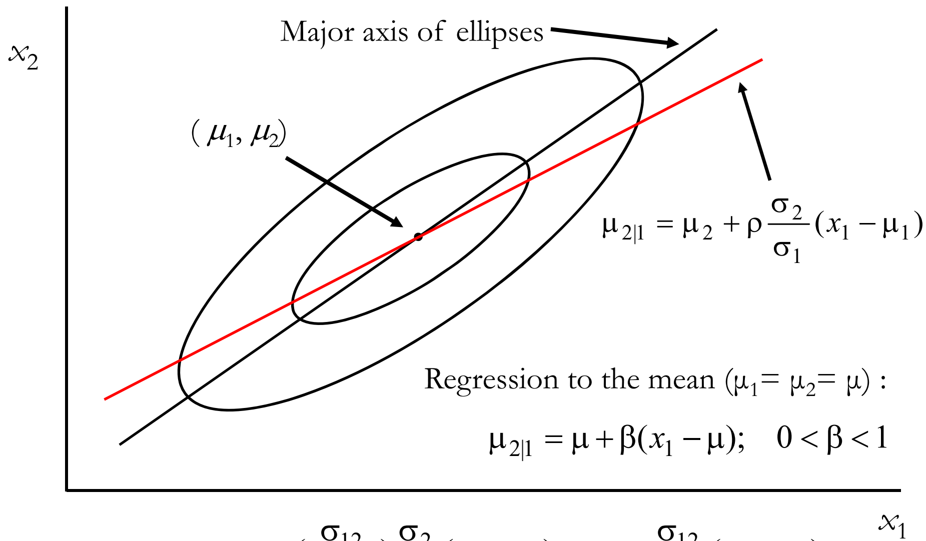
- Bivariate Normal Distribution with conditional distribution



- Using matrix notation, the conditional moments are given by:

$$\mu_{1|2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$$
$$\Sigma_{1|2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

Conditional distributions: Bivariate Normal



Note: $\mu_{2|1} = \mu_2 + (\frac{\sigma_{12}}{\sigma_1\sigma_2})\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1) = \mu_2 + \frac{\sigma_{12}}{\sigma_1^{\ 2}}(x_1 - \mu_1)$

Conditional distributions: KF Application

• The Kalman filter (KF) uses the observed data to learn about the unobservable state variables, which describe the state of the model.

• KF models dynamically what we measure, z_t, and the state, y_t. In the simple, linear model we have:

$y_t = \mathbf{A} y_{t-1} + w_t$ (*state or transition equation*)

$z_t = \mathbf{H} y_t + v_t$ (*measurement equation*)

w_t, v_t: error terms, with zero mean and variance $\mathbf{Q}$ and $\mathbf{R}$, respectively.

• Based on time t-1 information, the KF generates predictions for y_t:

$y_{t|t-1} = \mathbf{A}\, y_{t-1|t-1} + \mathbf{B}\, u_t$

$\mathbf{P}_{t|t-1} = \mathbf{A}\, \mathbf{P}_{t-1}\, \mathbf{A}^T + \mathbf{Q}$ (conditional variance of y_t)

• It also generates an update, once the information t is known:

$y_{t|t} = y_{t|t-1} + \mathbf{P}_{t|t-1}\, \mathbf{H}^T\, (\mathbf{F}_{t|t-1})^{-1}\, e_{t|t-1}$

$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\, \mathbf{H}^T\, (\mathbf{F}_{t|t-1})^{-1}\, \mathbf{H}\, \mathbf{P}_{t|t-1}$

28

Conditional distributions: KF Application

• We define the forecast error for the observed z_t and its variance as:

$$e_{t|t-1} = z_t - z_{t|t-1} = z_t - \mathbf{H} y_{t|t-1}$$
$$\mathbf{F}_{t|t-1} = \mathbf{H}\, \mathbf{P}_{t|t-1}\, \mathbf{H}^T + \mathbf{R}$$

Then, we write the joint of distribution of $(y_t, e_t)|I_t$:

$$\begin{pmatrix} y_t \mid I_{t-1} \\ e_t \mid I_{t-1} \end{pmatrix} \sim N\left(\begin{bmatrix} y_{t|t-1} \\ 0 \end{bmatrix}, \begin{bmatrix} P_{t|t-1} & P_{t|t-1}H' \\ HP_{t|t-1} & F_{t|t-1} \end{bmatrix} \right)$$

• Recall a property of the multivariate normal distribution:

$$\mu_{1|2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$$
$$\Sigma_{1|2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

Then, from the joint, we can easily derive the KF update:

$$y_{t|t} = y_{t|t-1} + \mathbf{P}_{t|t-1}\, \mathbf{H}^T\, (\mathbf{F}_{t|t-1})^{-1}\, e_{t|t-1}$$
$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\, \mathbf{H}^T\, (\mathbf{F}_{t|t-1})^{-1}\, \mathbf{H}\, \mathbf{P}_{t|t-1}$$