Avaliação de Desempenho

Distribuições Multivariadas Contínuas

Distribuição multivariada contínua

- A função de densidade de duas variáveis aleatórias X e Y conjuntamente distribuídas é uma função que mapeia o resultado de um experimento com duas variáveis aleatórias em um par ordenado em $\mathbf{R}^2$.

$$P[X \in A, Y \in B] = \int_B \int_A f_{X,Y}(x, y)\, dx\, dy$$

Exemplo 1

- Função de densidade conjunta de X e Y

$$f_{X,Y}(x,y)=\begin{cases} e^{-(x+y)} & \text{para } x \geq 0; y \geq 0 \\ 0 & \text{caso contrário} \end{cases}$$

- Validade da função

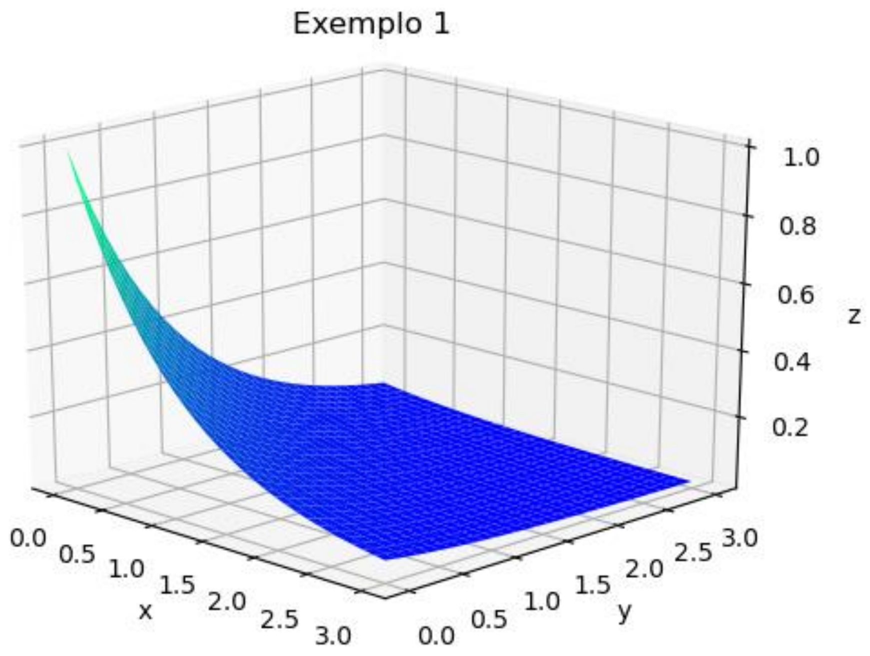
$$\int_0^\infty \int_0^\infty e^{-x} e^{-y} dx\, dy = 1$$

Integrate[exp(-(x+y))] {x from 0 to infinity} {y from 0 to infinity}

- Calcule $P[X>1, Y<1]$

$$P[X > 1, Y < 1] = \int_0^1 \int_1^\infty e^{-x} e^{-y} dx\, dy = 0{,}232544$$

Integrate[exp(-(x+y))] {x from 1 to infinity} {y from 0 to 1}



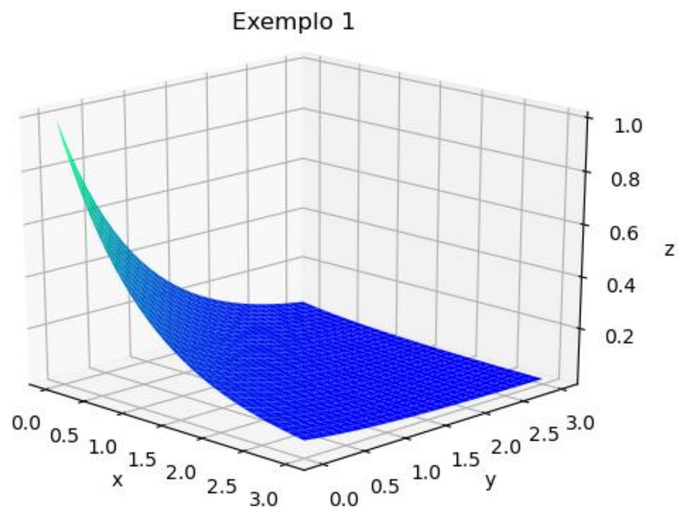
Função de densidade marginal

- Se a função de densidade conjunta $f_{X,Y}(x, y)$ é conhecida, é possível calcular $f_{\mathrm{X}}(x)$ e $f_{\mathrm{Y}}(y)$, ou seja, as funções de probabilidade das variáveis aleatórias X e Y.

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

$$\int_{-\infty}^{\infty} f_{X,Y}(x, y)dx = f_Y(y)$$

Exemplo 2



- Função de densidade conjunta de X e Y

$$f_{X,Y}(x,y) = \begin{cases} e^{-(x+y)} & \text{para } x \geq 0; y \geq 0 \\ 0 & \text{caso contrário} \end{cases}$$

Calcular $f_X(x)$ e $f_Y(y)$

$f_X(x) = e^{-x}$ Integrate[exp(-(x+y))] {y from 0 to infinite}

$f_Y(y) = e^{-y}$ Integrate[exp(-(x+y))] {x from 0 to infinite}

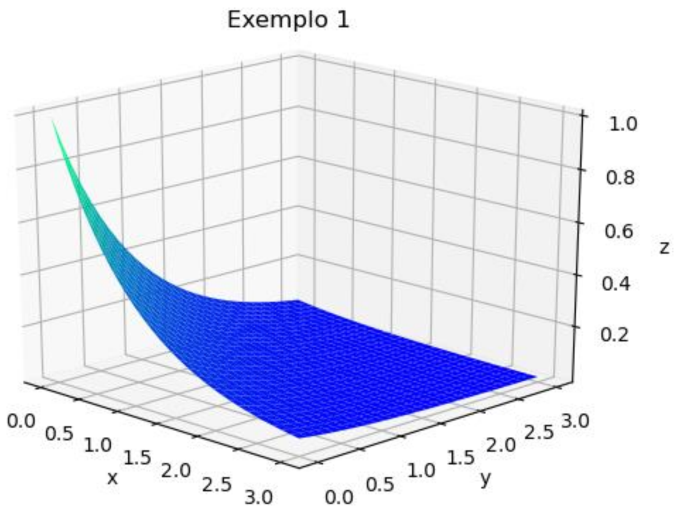
Independência

- Duas variáveis aleatórias X e Y são independentes se todos os eventos conjuntos são independentes, isto é, a ocorrência de um evento de uma variável aleatória não afeta a ocorrência de qualquer outro evento da outra variável
- X e Y são independentes se e somente se $f_{X,Y}(x, y) = f_X(xf_Y(y)$

Exemplo 3

- Função de densidade conjunta de X e Y

$$f_{X,Y}(x,y) = \begin{cases} e^{-(x+y)} & \text{para } x \geq 0; y \geq \\ 0 & \text{caso contrário} \end{cases}$$



$$f_X(x) = e^{-x} \qquad\qquad f_Y(y) = e^{-y}$$

$$f_X(x)\, f_Y(y) = e^{-x} e^{-y} = e^{-(x+y)} = f_{X,Y}(x,y)$$

Valor esperado

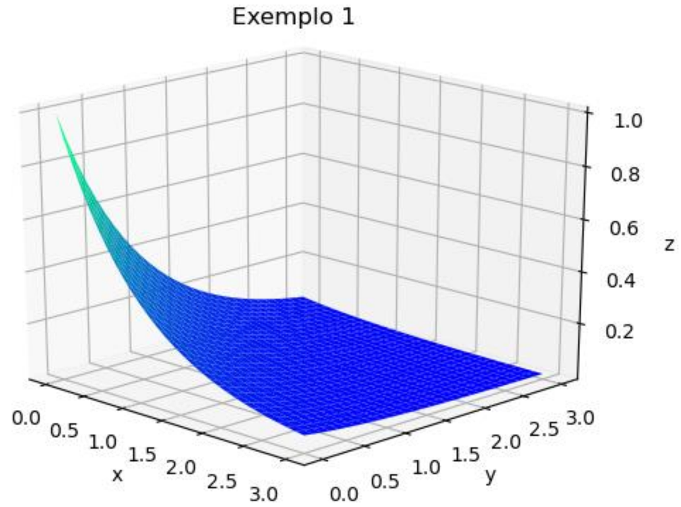
- O valor esperado de duas variáveis aleatórias contínuas conjuntamente distribuída é definido como o vetor composto pelo valor esperado de cada variável aleatória

$$E[X,Y]=\begin{bmatrix} E[X] \\ E[Y] \end{bmatrix}$$

Exemplo 4

- Função de densidade conjunta de X e Y

$$f_{X,Y}(x,y) = \begin{cases} e^{-(x+y)} & \text{para } x \geq 0; y \geq 0 \\ 0 & \text{caso contrário} \end{cases}$$



VALOR ESPERADO

$$f_X(x) = e^{-x} \qquad\qquad f_Y(y) = e^{-y}$$

$$E[X] = \int_0^{\infty} x f_X(x)\, dx = \int_0^{\infty} x e^{-x}\, dx = 1 \qquad E[Y] = 1$$

Integrate[exp(-x)] {x from 0 to infinite}

$$E[X,Y] = \begin{bmatrix} E[X] \\ E[Y] \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Função de distribuição acumulada conjunta

- A função de distribuição acumulada conjunta de duas variáveis aleatórias X e Y é definida por

$$F_{X,Y}(x, y) = P[X \le x, Y \le Y] = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{X,Y}(t,u)\,dt\,du$$

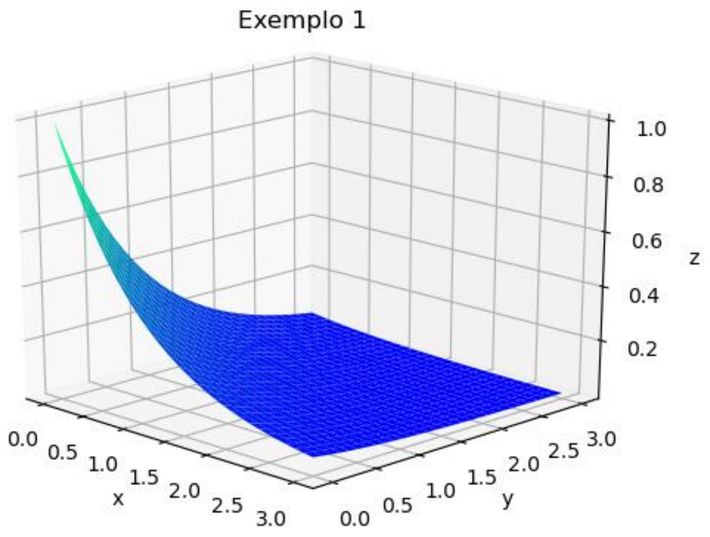
Exemplo 5

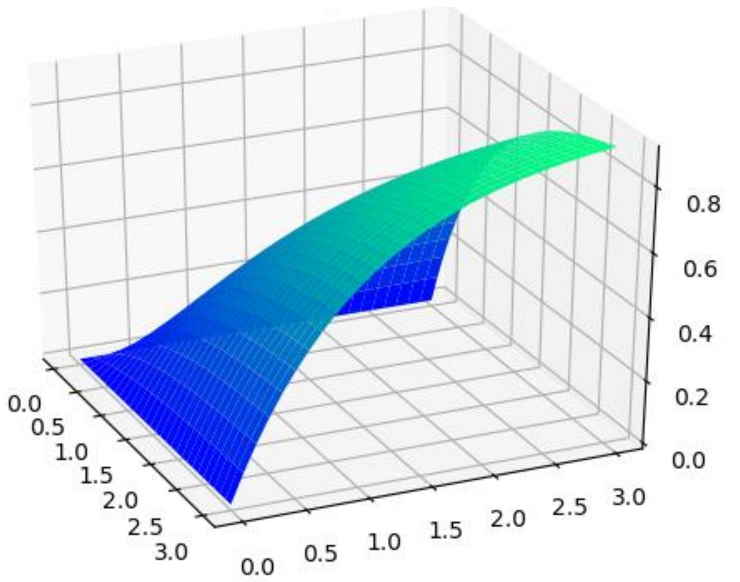
- *F*unção de densidade conjunta de X e Y

$$f_{X,Y}(x,y)=\begin{cases} e^{-(x+y)} & \text{para } x\geq 0; y\geq 0 \\ 0 & \text{caso contrário} \end{cases}$$

Calcular $F_{X,Y}(x, y)$

$$F_{X,Y}(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x} f_{X,Y}(t,u)dtdu=(1-e^{-x})(1-e^{-y})$$





Covariância

- Covariância é uma medida da relação de variação conjunta entre duas variáveis aleatórias (X e Y)
- Percebeu-se que a covariância entre duas variáveis depende do produto das duas variáveis e pode ser calculada da seguinte maneira

$$COV[X,Y] = E[X \cdot Y] - E[X] \cdot E[Y]$$

Exemplo 6

- Seja a seguinte distribuição conjunta de duas variáveis aleatórias discretas X e Y

Y \ X	0	1	2	$p_Y(y)$
1	3/20	3/20	2/20	8/20
2	1/20	1/20	2/20	4/20
3	4/20	1/20	3/20	8/20
$p_X(x)$	8/20	5/20	7/20	1

Exemplo 6

- Para calcular a covariância devemos calcular as esperanças de X e Y:

$$E[X] = 0 \times \frac{8}{20} + 1 \times \frac{5}{20} + 2 \times \frac{7}{20} = \frac{19}{20}$$

$$E[Y] = 1 \times \frac{8}{20} + 2 \times \frac{4}{20} + 3 \times \frac{8}{20} = 2$$

Exemplo 1

$$COV[X,Y]=(0-\frac{19}{20})\times(1-2)\times\frac{3}{20}+(1-\frac{19}{20})\times(1-2)\times\frac{3}{20}+(2-\frac{19}{20})\times(1-2)\times\frac{2}{20}+$$

$$+(0-\frac{19}{20})\times(2-2)\times\frac{1}{20}+(1-\frac{19}{20})\times(2-2)\times\frac{1}{20}+(2-\frac{19}{20})\times(2-2)\times\frac{2}{20}+$$

$$+(0-\frac{19}{20})\times(3-2)\times\frac{4}{20}+(1-\frac{19}{20})\times(3-2)\times\frac{1}{20}+(2-\frac{19}{20})\times(3-2)\times\frac{3}{20}=0$$

Covariância

- O sinal na covariância indica a direção da relação de variação conjunta entre as duas variáveis
  - Um sinal positivo indica que elas movem juntas e um negativo que elas movem em direções opostas
  - A covariância cresce com a força do relacionamento, mas é relativamente difícil fazer julgamentos sobre o força do relacionamento entre as duas variáveis observando a covariância porque ela não é uma medida padronizada
- É possível provar que

$$-DP[X] \cdot DP[Y] \leq COV[X,Y] \leq +DP[X] \cdot DP[Y]$$

Correlação

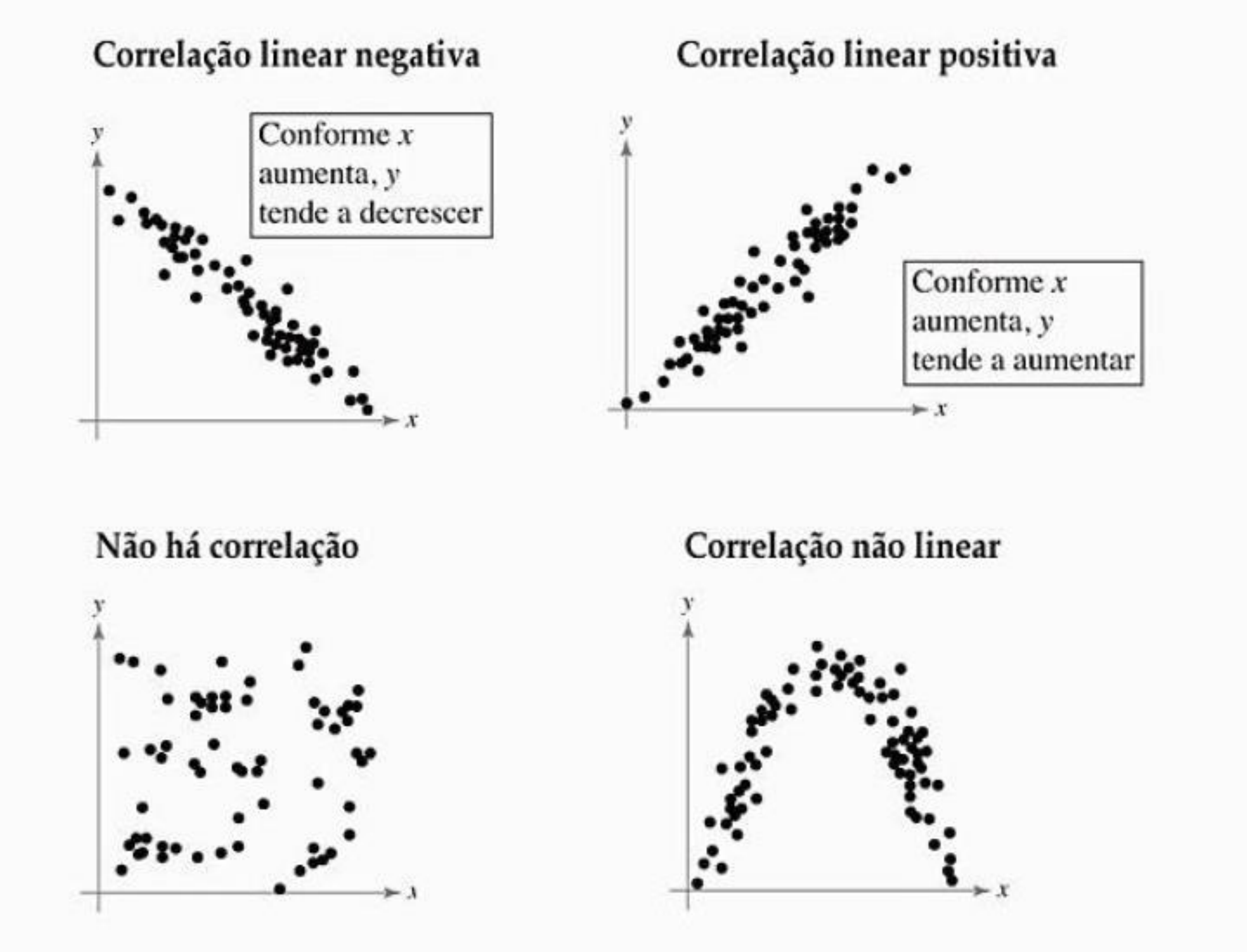
- A correlação é a medida padronizada da relação entre duas variáveis
- É calculada a partir da covariância

$$COR[X,Y] = \frac{COV[X,Y]}{DP[X] \cdot DP[Y]}$$

Correlação

- A correlação nunca pode ser maior do que 1 ou menor do que -1
- Uma correlação próxima a zero indica que as duas variáveis não estão linearmente relacionadas
- Uma correlação positiva indica que as duas variáveis movem juntas, e a relação é forte quanto mais a correlação se aproxima de um
- Uma correlação negativa indica que as duas variáveis movem-se em direções opostas, e que a relação também fica mais forte quanto mais próxima de -1 a correlação ficar
- Duas variáveis que estão perfeitamente correlacionadas positivamente ($COR[X,Y]=1$) movem-se em perfeita proporção na mesma direção
- Duas variáveis que estão perfeitamente correlacionados negativamente movem-se em perfeita proporção em direções opostas

Correlação


Correlação linear negativa
y
Conforme x aumenta, y tende a decrescer
x
Correlação linear positiva
y
Conforme x aumenta, y tende a aumentar
x
Não há correlação
y
x
Correlação não linear
y
x

Matriz de covariância

- Se os elementos de um vetor coluna

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \cdots \\ X_n \end{bmatrix}$$

  forem variáveis aleatórias (cada uma com variância finita), então a matriz de covariância será a matriz cujo elemento (i, j) é $COV(X_i,X_j)$

$$\sum = \begin{bmatrix} COV[X_1, X_1] & COV[X_1, X_2] & \cdots & COV[X_1, X_n] \\ COV[X_2, X_1] & COV[X_2, X_2] & \cdots & COV[X_2, X_n] \\ \vdots & \vdots & \ddots & \vdots \\ COV[X_n, X_1] & COV[X_n, X_2] & \cdots & COV[X_n, X_n] \end{bmatrix}$$

- A covariância entre um elemento X_i e ele mesmo é a sua variância e forma a diagonal principal da matriz

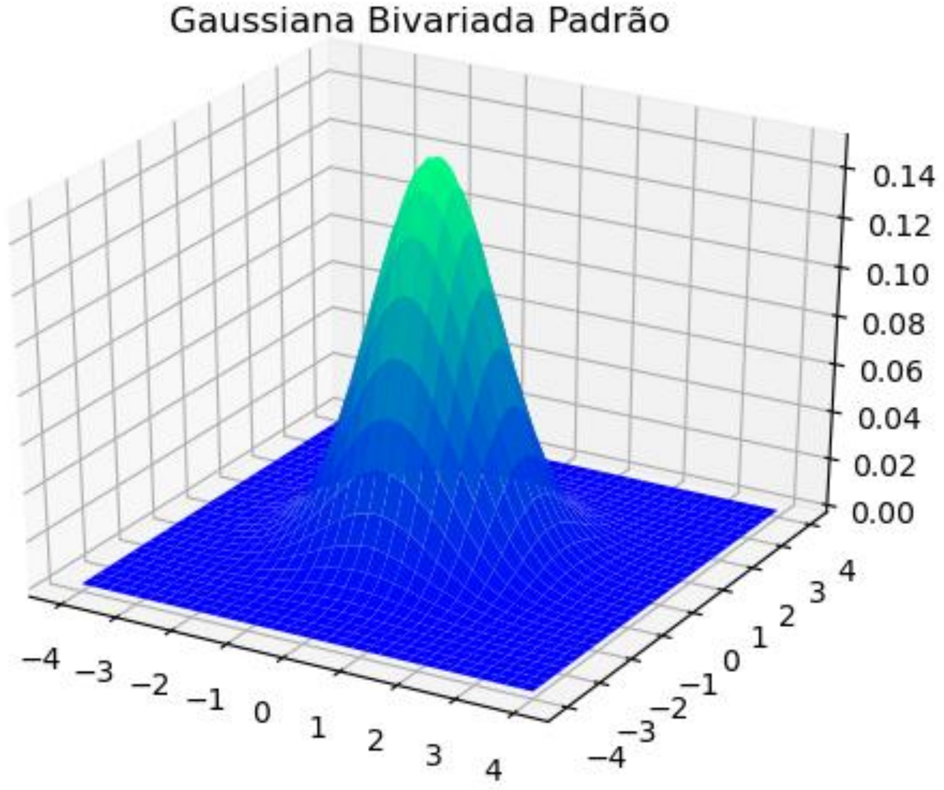
Gaussiana bivariada padrão

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}(x^2 + 2\rho xy + y^2)\right]$$

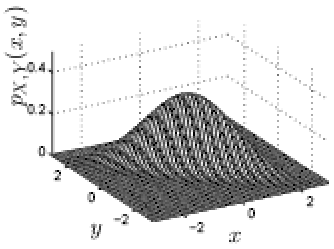
-∞ < x < ∞; -∞ <y <∞

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}x^2\right] \qquad f_Y(y) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}y^2\right]$$

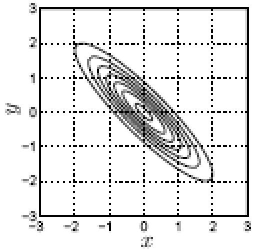
$$COV[X,Y] = \rho$$



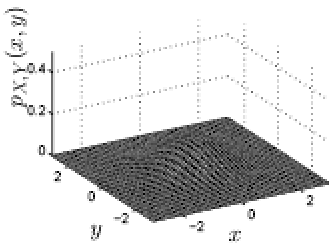
Gaussiana bivariada padrão



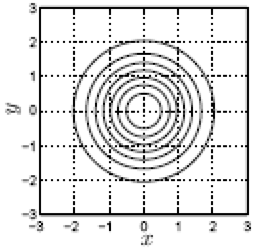
(a) $\rho = -0.9$



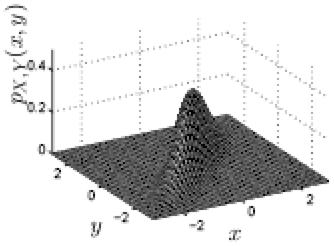
(b) $\rho = -0.9$



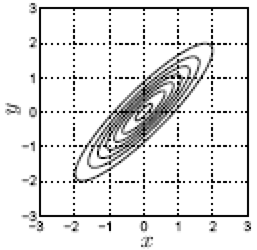
(c) $\rho = 0$



(d) $\rho = 0$



(e) $\rho = 0.9$



(f) $\rho = 0.9$

Gaussiana multivariada

- Formada pela distribuição conjunta de n variáveis aleatórias gaussianas
- Função de densidade de probabilidade

$$f_X(\mathbf{X}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{\Sigma}|}} \exp\left(-\frac{1}{2} (\mathbf{X}-\boldsymbol{\mu})^{\mathrm{T}} \mathbf{\Sigma}^{-1} (\mathbf{X}-\boldsymbol{\mu}) \right)$$

- Dois parâmetros

$$\boldsymbol{\mu} = \begin{bmatrix} E[X_1] \\ E[X_2] \\ \cdots \\ E[X_n] \end{bmatrix} \qquad \Sigma = \begin{bmatrix} COV[X_1, X_1] & COV[X_1, X_2] & \cdots & COV[X_1, X_n] \\ COV[X_2, X_1] & COV[X_2, X_2] & \cdots & COV[X_2, X_n] \\ \vdots & \vdots & \ddots & \vdots \\ COV[X_n, X_1] & COV[X_n, X_2] & \cdots & COV[X_n, X_n] \end{bmatrix}$$

$|\Sigma|$ é o determinante de Σ $\qquad$ Σ^{-1} é a inversa de Σ

Avaliação de Desempenho

Distribuições Multivariadas Contínuas