

Polinomios y Estadística

Universidad de Concepción, Chile
Departamento de Geofísica
Programación Científica con Software libre

Primavera, 2011



Contenidos

- 1 Funciones polinomiales
- 2 Estadística descriptiva
- 3 Utilidades

Polinomios

Un polinomio puede ser representado por sus coeficientes, por ejemplo

$$p(x) = x^3 - 2x^2 - 3x + 2$$

es caracterizado por el vector $p = [1, -2, -3, 2]$

Consideraciones :

- Los coeficientes deben ser ordenados en forma decreciente por su grado
- Deben ser representados todos los coeficientes aun si su valor es 0

Polinomios

Una vez representado el polinomio se pueden realizar varias operaciones asociadas a este, entre algunas :

- Las raíces del polinomio son estimadas usando el comando **roots**
- La función **polyval** se utiliza para evaluar el polinomio en un dominio dado
- La función **polyfit** ajusta un set de datos a un polinomio

Polinomios

Ejemplo 1

Encontrar las raíces del polinomio

$$2x^4 - 7x^3 - 8x^2 + 14x + 8$$

Comprobemos el resultado obtenido con **roots** usando la función **polyval** y graficando.

y = polyval(p,x) entrega los valores del polinomio p evaluado en x, donde x puede ser vector o escalar.

Solución

```
# solucion  
  
p = [2 -7 -8 14 8];  
c = roots(p); # raices  
x = [-10:.1:10]; # dominio  
y = polyval(p,x); # evaluamos polynomio  
# graficamos  
plot(x,y)  
# zoom para ver raices  
axis([-10 10 -50 50])
```

Polinomios

Ejemplo 2

p = polyfit(x,y,n) entrega los coeficientes del polinomio de grado n que mejor ajusta los datos $p(x(i))$ a $y(i)$ en el sentido de mínimos cuadrados.

Descargue el archivo **fit.dat** y ajuste los datos a un polinomio de grado 2.

Solución

```
# ajuste por minimos cuadrados  
  
d = load('fit.dat');  
x = d(:,1);  
y = d(:,2);  
  
# buscamos los coeficientes del polinomio  
# de grado 2  
p = polyfit(x,y,2);  
  
# comportamiento de el polinomio  
# ajustado en relacion a los datos  
y1 = polyval(p,x);  
  
plot(x,y,'o',x,y1,'r')  
grid
```


Estadística

Medidas de tendencia central

Media aritmética

Se puede definir como el centro de gravedad de una distribución, el cual no está necesariamente en la mitad. Su representación matemática es

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

En señales geofísicas es común restarle la media a la señal y trabajar solo con la anomalía en torno a este valor medio, de alguna forma esto permite menos propagación de errores. Así tenemos que la nueva señal es

$$\hat{x} = x - \bar{x}$$

En octave tenemos la función **mean** para calcular la media de un array, en el caso de una matriz la media es calculada sobre cada columna.

Descargue el archivo **unidata.dat** y muevalo a su directorio de trabajo, luego importe los datos usando la función **load**. Calcule la media de todas las columnas que componen la matriz, excepto la primera.

Solución

```
data = load('unidata.dat');  
mean = mean(data(:,2:end));
```

Estadística

Medidas de dispersión

Desviación estándar

Dispersión de los datos en torno al valor medio. Para distribuciones no conocidas la desviación estándar muestral se calcula como

$$s = \sqrt{\frac{\sum_{i=1}^N (\bar{x} - x_i)^2}{N - 1}}$$

En octave la función **std** se utiliza para calcular la desviación estándar de un array, en el caso de una matriz la desviación estándar es calculada sobre cada columna.

Para el archivo **unidata.dat** realice el cálculo anterior, pero esta vez calcule la desviación estándar.

Solución

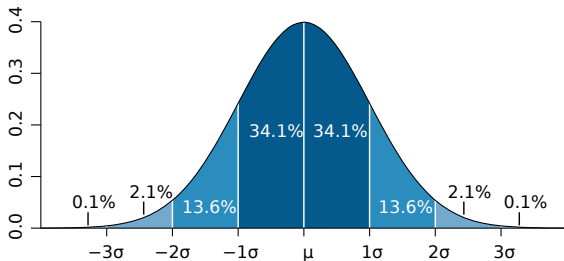
```
data = load('unidata.dat');  
sigma = std(data(:,2:end));
```

Probabilidades

Funciones de distribución

La *distribución de probabilidad* de una variable aleatoria es una función que asigna a cada suceso definido sobre la variable aleatoria la probabilidad de que dicho suceso ocurra.

Fig: Distribución Gaussiana



FDP & CDF

Entre las funciones de densidad de probabilidad (FDP) que incorpora octave encontramos las siguientes

- **normpdf** normal
- **tpdf** t-student
- **chi2pdf** chi-cuadrado
- **unifpdf** uniforme

y para las funciones de distribución de probabilidad

- **normcdf** normal
- **tcdf** t-student
- **chi2cdf** chi-cuadrado
- **unifcdf** uniforme

Ejemplo

`cu = normcdf(x, μ , σ)`

entrega la probabilidad de observar un valor menor que x , en un experimento donde las variables aleatorias se distribuyen normal con media μ y desviación estándar σ

Ejercicio

La cantidad de agua caída sobre una ciudad se distribuye normal, si la media del agua precipitada son 100 mm y la desviación típica 20 mm .¿ Cual es la probabilidad que el agua caída sea superior a 80 mm ?

Funciones de conversión

num2str

```
str = num2str(A, format)
```

convierte el array **A** a su representación *string*. La palabra **string** se usa para hacer referencia a una cadena de caracteres. Entre los atributos que despliega **whos** se encuentra **Class**, que proporciona información referente a la *clase de variable*. De esta manera con el comando **whos** podemos constatar que si declaramos un string, por ejemplo si `dia='lunes'`, este será de clase **char**, relativo a *character*.

En **format**, entrecomilla simple, las opciones de conversión. Algunas de estas opciones son listadas en la siguiente tabla.

Carácter de conversión	Descripción
%c	Secuencia de caracteres.El número esta especificado por el ancho del campo, ej: %10c lee 10 caracteres.
%d	Enteros en base decimal.
%f	Números en punto flotante.
%s	Palabras.Secuencia de caracteres hasta antes de un espacio en blanco.
%e	Notación exponencial, tal como 3.141593e+00
%i	Número entero

Funciones de conversión

datenum

`fecha = datenum(A, M, D)`

Retorna una fecha numérica para los correspondientes elementos de A, M , D (año, mes, día).

Ejemplo

Crear un vector numérico de fechas entre el 23 de Agosto de 1982 al 23 de Febrero de 2010.

```
vi = datenum(1982,08,23);  
vf = datenum(2010,02,23);  
vc = vi:vf;
```

Funciones de conversión

datetick

`datetick(eje, tipo)`

Cambia el tipo de etiquetado del eje x, y, z a un formato definido por *tipo*. Para un correcto resultado, los valores del eje especificado deben estar en formato de fecha numérico (datenum).

Ejemplo

Poner etiquetado al eje X de tipo día/mes/año

```
datetick('x',2)
```