Esercizio 1

Sia (X_1,X_2, …X_n) un campione casuale di VA X di tipo esponenziale ($\lambda e^{-\lambda x}$). Il parametro λ è incognito. Si determini lo stimatore basato sulla massima verosimiglianza per il parametro λ.

La Pdf è $\lambda e^{-\lambda x}$ per ogni Xi. Si ipotizza che gli esperimenti siano indipendenti, quindi la PDF congiunta è il prodotto delle marginali. La funzione verosimiglianza è la congiunta, e va massimizzata in modo da trovare il parametro λ.

$$L(\lambda)=\prod_{i=1}^{n}\lambda e^{-\lambda x_i}$$

Conviene trasformare logaritmicamente ambo i membri

$$\ln L(\lambda)=\ln\prod_{i=1}^{n}\lambda e^{-\lambda x_i}=\sum_{i=1}^{n}\ln\lambda+\sum_{i=1}^{n}\ln(e^{-\lambda x_i})=n\ln\lambda-\lambda\sum_{i=1}^{n}x_i$$

il massimo si determina derivando rispetto a λ, uguagliando a zero e risolvendo per λ:

$$\frac{d}{d\lambda}\ln L(\lambda)=\frac{n}{\lambda}-\sum_{i=1}^{n}x_i=0$$

$$\lambda=\frac{n}{\sum_{i=1}^{n}x_i}$$

Esercizio 2

Misure di concentrazione di una specie chimica (c) all'interno di un reattore al variare della pressione di esercizio (P) sono di seguito riportate:

P	0.51	0.56	0.63	0.71	0.80	0.86	0.93	0.99
c	4.25	4.47	4.70	4.89	5.41	5.73	6.46	7.01

Si suppone che un buon modello per descrivere i dati sia:

$$c(P)=3^{A\cdot\exp(P)+B\cdot P^2+\frac{C}{P}}$$

dove A, B, e C sono parametri del modello.

a) si effettui una stima dei parametri del modello linearizzato
b) si effettui una stima dei parametri del modello non lineare
c) si effettui una stima della varianza del modello non lineare attraverso il metodo della massima verosimiglianza
d) si diagrammino i dati sperimentali e la previsione del modello non lineare

Il modello può essere linearizzato attraverso la trasformazione logaritmo naturale:

$$\ln c(P) = \ln 3^{A\cdot \exp(P)+B\cdot P^2+\frac{C}{P}} = \ln 3\left[A\cdot \exp(P)+B\cdot P^2+\frac{C}{P}\right] \qquad (1)$$

oppure applicando la trasformazione $\log_3$:

$$\log_3 c(P) = \log_3 3^{A\cdot \exp(P)+B\cdot P^2+\frac{C}{P}} = A\cdot \exp(P)+B\cdot P^2+\frac{C}{P} \qquad (2)$$

Ricordando che: $\log_3 x = \ln x / \ln 3$, il primo membro può essere riscritto come:

$$\log_3 c(P) = \ln c(P) / \ln(3)$$

l'equazione (2) si riduce alla (1).

Considerando la (1), la seguente function da utilizzare con il comando *fminsearch* fornisce una stima dei parametri del modello linearizzato:

```
function f = Lin(c)
x = [0.51, 0.56, 0.63, 0.71, 0.80, 0.86, 0.93, 0.99];
y = [4.25, 4.47, 4.70, 4.89, 5.41, 5.73, 6.46, 7.01];
a1 = c(1);
a2 = c(2);
a3 = c(3);
expP = exp(x)*log(3);
P2 = x.^2*log(3);
invP = 1./x*log(3);
clin = log(y);
w = 1./(y.^2);
f = sum(((clin - a1*expP - a2*P2 - a3*invP).^2)./w);
```

Si noti la presenza dei pesi *w* e l'inclusione del fattore *ln 3* nella definizione della variabile dipendente del modello linearizzato.

I parametri stimati sono:

$$\hat{A}_{lin} = 0.5842$$
$$\hat{B}_{lin} = 0.0170$$
$$\hat{C}_{lin} = 0.1812$$

Utilizzando tali stime come primo tentativo nel comando *nlinfit* si perviene alla seguente stima dei parametri del modello non lineare:

$$\hat{A}_{lin} = 0.5840$$
$$\hat{B}_{lin} = 0.0176$$
$$\hat{C}_{lin} = 0.1811$$

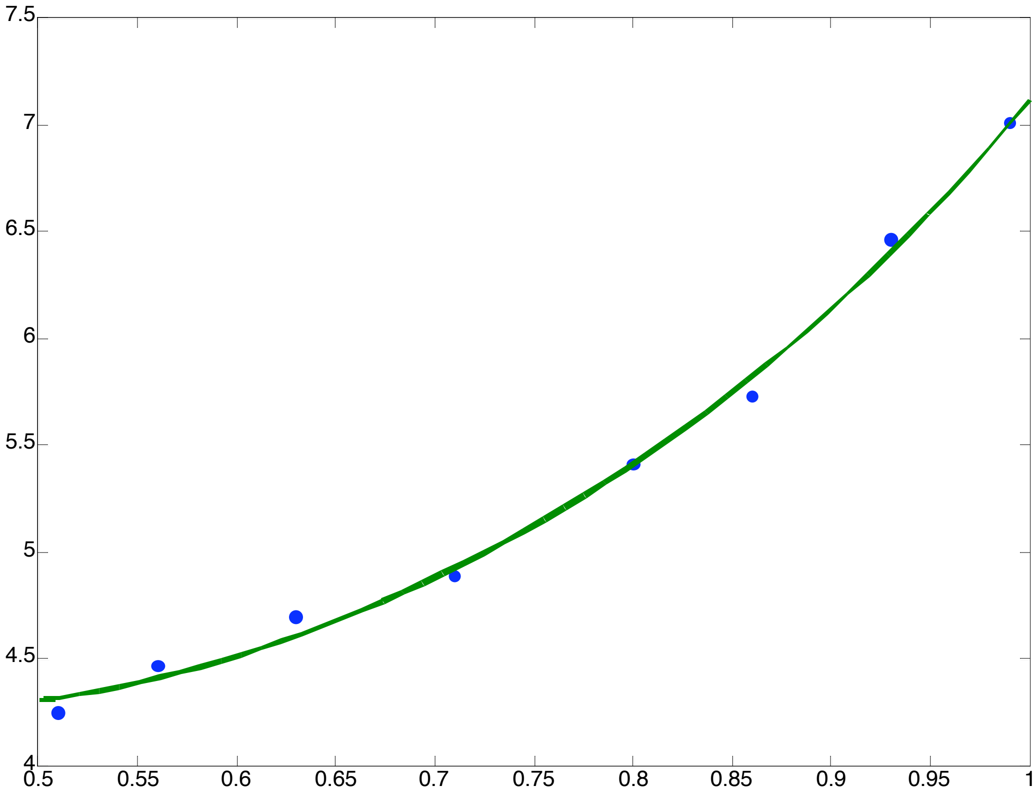
La varianza sperimentale (imparziale) può essere stimata con la formula:

$$\hat{\sigma}^2 = \frac{1}{N-3}\sum_{i=1}^{N}(y_i - f(x_i,\boldsymbol{\theta}))^2 \qquad (3)$$

dove N = 8 ed f è il modello non lineare.

La formula (3) fornisce $\hat{\sigma}^2 = 0.006$.

Infine, il diagramma dei dati sperimentali e del modello è:



Esercizio 3

In basso sono riportati si risultati di una sperimentazione con esperimenti indipendenti. Due modelli alternativi sono proposti per descrivere il processo: 1) $a + bx$; 2) $a + bx^2$

i) Si proceda alla stima dei parametri a e b per entrambi i modelli ipotizzando che i risultati sono stati ottenuti con esperimenti indipendenti caratterizzati da un errore di misura gaussiano con media zero e varianza indipendente dalle condizioni di misura.
ii) Si calcoli la varianza sperimentale per entrambi i casi.
iii) Si esegua la diagnostica sui risultati ottenuti utilizzando i residui analizzando criticamente i risultati in relazione alle ipotesi fatte.
iv) Per il miglior modello si determinino gli intervalli di fiducia dei parametri.
v) Se la diagnostica per il miglior modello risultasse comunque insoddisfacente, si provi a suggerire possibili miglioramenti della stima

x	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.
y	0.29	0.48	0.96	1.2	1.6	2.3	3.5	3.2	4.7	5.7	6.4	8.2	11.	12.	13.	14.	17.	18.	19.

Con Matlab si possono determinare i parametri dei due modelli lineari ottenendo

Modello 1: a=-4.61, b=2.2

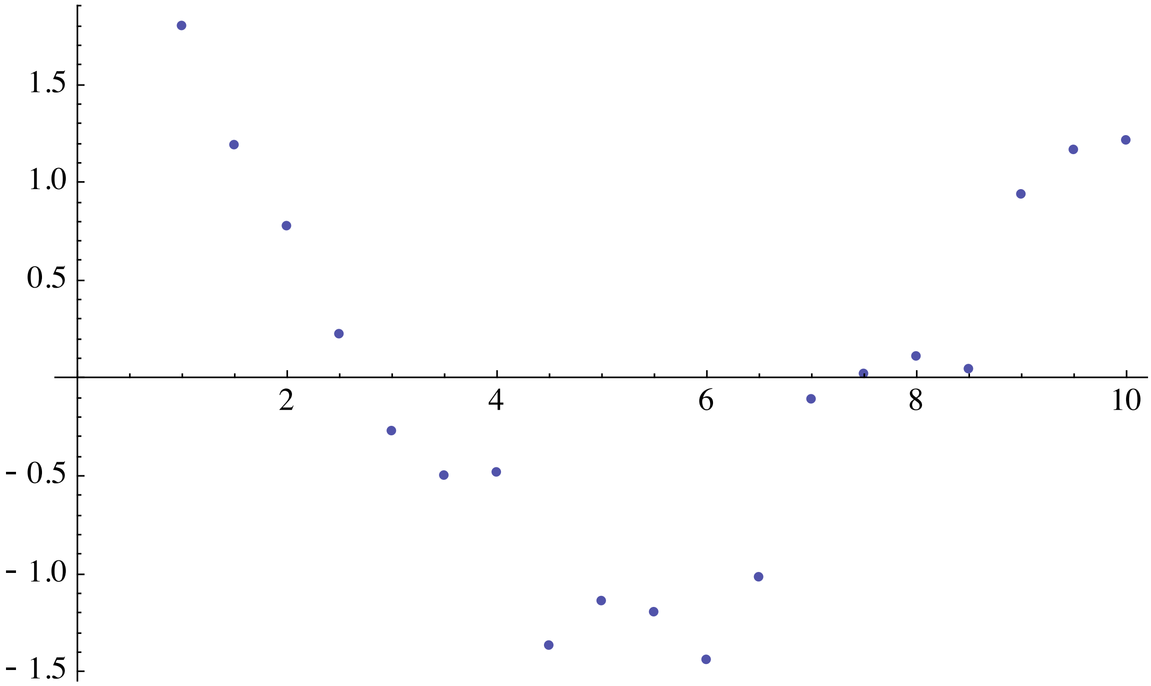
Modello 2: a=-0.055, b=0.20

La varianza si determina con lo stimatore imparziale: $s^2 = \dfrac{(y - \mathbf{X} \cdot \boldsymbol{\theta})' . (y - \mathbf{X} \cdot \boldsymbol{\theta})}{n - 2}$, si ottiene

Modello 1: s^2=2.32;   Modello 2: s^2=0.335

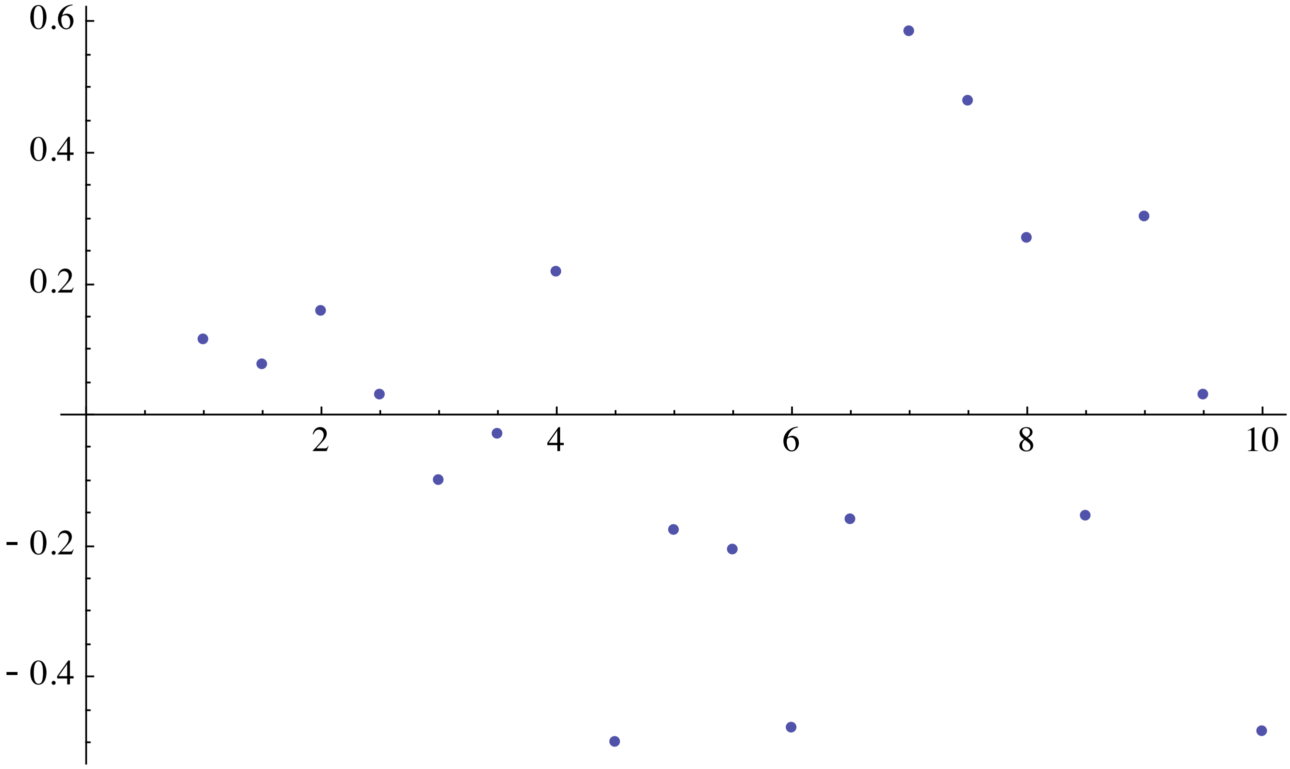
Si può già notare una minore varianza nel caso del modello 2.

Diagramma dei residui standardizzati contro la condizione sperimentale
Modello 1



Si nota una dipendenza dei residui dalle condizioni sperimentali che suggerisce una non adeguatezza del modello.

Modello 2



In questo caso i residui sono distribuiti attorno allo zero, ma si nota una struttura che suggerisce la presenza di eteroschedasticità. Probabilmente l'ipotesi di varianza sperimentale indipendente dalla condizione sperimentale non era valida, ma il modello può comunque ben descrivere i dati.

Si suggerisce di modellare l'esperimento per tenere conto della eteroschedasticità.
