

Report su Case study del gruppo Lubjana G21

Borqal Amin 1073928, Iacoban Loris 1074130, Moressa Andrea 1074124

January 14, 2022

1 Introduzione

Abbiamo sviluppato un programma in MatLab attraverso il quale con delle tecniche di regressione lineare multipla abbiamo ottenuto dei modelli che descrivono la relazione di un gruppo di variabili esplicative riguardanti la meteorologia ed il consumo dei carburanti con gli inquinanti NO2 e PM10. Il programma MatLab utilizza i dati di una tabella proveniente dal sito dell'ARPA Lombardia. Il nostro studio consiste nel cercare capire quale relazione intercorre tra gli inquinanti PM10 ed NO2 con i 2 pacchetti di dati riguardanti la meteorologia ed il consumo dei carburanti. Abbiamo utilizzato il metodo dello stepwise, un algoritmo che automaticamente rimuove (o aggiunge) una variabile alla volta al modello di regressione. Il modello migliore è quindi scelto in base alla significatività dei vari coefficienti di regressione.

2 Ricerca dei 2 modelli migliori per PM10 e NO2

2.1 Modelli di regressione per il PM10

Prendiamo l'inquinante PM10 come variabile dipendente e studiamo la sua relazione con i dati riguardanti la meteorologia prima e con il consumo dei carburanti dopo. Inanzitutto occorre verificare che il determinante della matrice $X'X$ sia maggiore di zero. per prima cosa si aggiunge una colonna di uni e poi calcolare il determinante con la funzione "det" fornita da matlab. I determinanti risultano entrambi maggiori di zero perciò vale la condizione di esistenza. Nel caso non lo fossero stati non sarebbe stato possibile ottenere dei buoni modelli.

2.2 Confronto tra i 2 modelli

Una volta ottenuti i 2 modelli cerchiamo di capire quale sia il migliore dei 2 mettendo a confronto i coefficienti elencati in seguito.

```
R2PM10=ModelloPm10Carb.Rsquared.Adjusted
R2PM101=ModelloPm10Temp.Rsquared.Adjusted
%Notiamo che il modello PM10Temp ha r maggiore
PVPM10=ModelloPm10Carb.Coefficients.pValue
PVPM101=ModelloPm10Temp.Coefficients.pValue
%i p-value sono prossimi allo zero e rifiutano H0
FtestPM10=coefTest (ModelloPm10Carb)
```

```

FltestPM10=coefTest (ModelloPm10Temp)
%Tutte e 4 molto piccoli tendenti a zero
%-----PM10Carb-----
MSEPM10=ModelloPm10Carb.MSE %abbastanza significativo dato che vale
    122.4591
ResiduiPM10 = ModelloPm10Carb.Residuals.Raw
Media=mean(ResiduiPM10) %Media dei residui bassa
plot (ResiduiPM10)
ylines(0,'r','LineWidth',3)
ylines(mean(ResiduiPM10),'b','LineWidth',2)
histfit (ResiduiPM10)
%-----PM10Temp-----
MSEPM101=ModelloPm10Temp.MSE % 48.3615 minore del modello Carb
ResiduiPM101 = ModelloPm10Temp.Residuals.Raw
Media1=mean(ResiduiPM101) %Media dei residui bassa

```

L'Rquadrato del modello PM10-Meteorologia è più alto di quello ottenuto dal suo "competitor".Questo significa che la deviazione dei punti effettivi dai punti adattati, in media, è più piccola rispetto a quella del secondo modello. pvalue e analisi dei residui ci portano alla conclusione che il modello PM10-Meteorologia trovato è il migliore.

2.3 Ricerca del miglior modello per l'NO2

Prendiamo l'inquinante NO2 come variabile dipendente e studiamo la sua relazione con i dati riguardanti la meteorologia prima e con il consumo dei carburanti dopo. Ommettiamo il codice MatLab e saltiamo subito alla conclusione: I determinanti delle rispettive matrici XX' sono maggiori di zero e analizzando quale modello sia migliore tra i 2 in base ai coefficienti citati prima e all'analisi dei residui scopriamo che anche qui il modello NO2-Meteorologia è migliore del NO2-ConsumoCarburanti.

3 Conclusione

Analizzando i vari modelli per l'inquinante PM10 abbiamo riscontrato che il modello con i risultati migliori è il modello PM10 Meteorologia ,invece per l'inquinante NO2 il modello NO2-meteorologia. Possiamo vedere come vi è una maggiore relazione tra il PM10 e le statistiche meteorologiche in confronto all'NO2.