Lista 2

Aneta Przydróżna

Regresja grzbietowa

Generuję macierz $X_{1000x950}$, tak aby X'X=I, wektor predyktorów z k pierwszymi wyrazami równymi 3, resztą równą 0 oraz wektor szumu ϵ z rozkładu $N(0, I)$.

• Macież X jest ortonormalna, zatem optymalna $\lambda = \frac{950 \cdot \sigma^2}{\|\beta\|^2}$ najlepiej minimalizuje $MSE(\hat{\beta})$. Poniżej znajdują się λ dla k={20,100,200} oraz wybrane statystyki policzone z użyciem tego estymatora.

```
##      lambda Obciazenie Wariancja     MSE  moc
## 20   5.278    127.222    24.105 151.327 0.20
## 100  1.056    237.327   224.836 462.162 0.18
## 200  0.528    214.810   407.008 621.818 0.19
```

Z kolejnymi istotnymi zmiennymi objaśniającymi, optymalna lambda jest coraz mniejsza.

- $Bias^2(\hat{\beta}) = \frac{\lambda^2 \cdot \|\hat{\beta}\|^2}{(1+\lambda)^2}$
- $Var(\hat{\beta}) = \frac{\sigma^2 \cdot 950}{(1+\lambda)^2}$
- $MSE(\hat{\beta})$=$Bias^2(\hat{\beta})$+$Var(\hat{\beta})$
- Wartość krytyczna $C = \frac{\sigma}{1+\lambda} \cdot \Phi^{-1}(1 - \frac{\alpha}{2 \cdot 950})$

Używamy tej wartości aby testować hipotezę zerową, że β_i=0, dla i={1,..,950}. Gdy $|\hat{\beta_i^R}| > C$, odrzucamy H_0. Wskaźnik błędu $FWER \leq \alpha = 0.1$ to prawdopodobieństwo popełnienia jednego lub większej liczby fałszywych odkryć, w tym przypadku na poziomie istotności $\alpha = 0.1$.

Podana moc wskazuje stosunek liczby prawdziwych odkryć, a niezerowych parametrów. We wszystkich przypadkach znajduje się bliżej zera niż jedynki, co świadczy o nie za dobrym działaniu wskaźnika. Problem związany z poprawką Bonferroniego jest taki, ze ograniczając liczbą fałszywych pozytywów przez obnizenie progu p-wartości przyjmujemy niepoprawne hipotezy zerowe, a więc nie wykrywamy istotnych cech.

Dla metody najmniejszych kwadratów mamy następujące parametry:

- Estymator β jest nieobciążony zatem $Bias^2(\hat{\beta}) = 0$
- $Cov(\hat{\beta}) = \sigma^2 \cdot I = I$
- $Var(\hat{\beta}) = MSE(\hat{\beta}) = Tr(Cov(\hat{\beta})) = 950$

Empiryczne wyniki dla powyższych parametrów:

```
## [1] "Ridge"
```

```
##     Obciążenie Wariancja     MSE   Moc
## 20     127.410    23.600 151.442 0.187
## 100    238.567   220.121 463.271 0.196
```

```
## 200    219.403   398.473 619.818 0.191
```

```
## [1] "OLS"
```

```
##     Obciążenie Wariancja      MSE   Moc
## 20       4.592   947.430 952.3940 0.193
## 100      4.623   946.983 945.2341 0.185
## 200      4.935   949.324 953.5630 0.201
```

Wyliczone teoretycznie i empirycznie dane w obu przypadkach są bardzo podobne, natomiast różnią się między sobą. Korzystając z metody najmniejszych kwadratów powinniśmy spodziewać się zerowego obciążenia, natomiast pojawiło się ono, ze względu na liczbę eksperymentów, lecz bardzo niewielkie. W przypadku regresji grzbietowej dla dużych k, wartości obciążenia przekraczają nawet 200. Kosztem tego metoda osiąga znacznie lepsze wariancje i MSE. Najlepiej widać to w przypadku, gdy k=20, gdy wariancja Ridge dla wektora estymatora β jest niemalże 40-krotnie mniejsza, MSE 6-krotnie. Ridge reaguje również na ilość istotnych parametrów, w OLS wszystkie parametry są zbliżone, niezależnie od k. Ogólnie można swtwierdzić, że metoda regresji grzbietowej radzi sobie lepiej niż metoda najmniejszych kwadratów.