Lecture 1

Jussi Klemelä

January 23, 2009

1 Introduction

The preliminary schedule.

1. Regression function estimation
   Portfolio selection.
2. Local averaging: kernel regression
3. Local averaging: nearest neighborhood regression
4. Estimation of conditional variance and conditional quantiles
   Option pricing
5. Empirical risk minimization
6. Empirical risk minimization continued
7. Support vector machines
8. Tree based methods
9. Aggregation
10. Linear models, generalized linear models, single index models
11. Local likelihood
12. Kertaus

2 Regression function estimation

We observe a sequence

$$(Y_1, X_1), \ldots, (Y_n, X_n)$$

of random variables, where $Y_i \in \mathbf{R}$ and $X_i \in \mathbf{R}^d$. The variables $Y_1, \ldots, Y_n$ are called the response variables and variables $X_1, \ldots, X_n$ are called the explanatory variables. In the classical regression setting we assume that $(Y_1, X_1), \ldots, (Y_n, X_n)$ is an identically distributed sequence. We want to estimate the conditional expectation $f : \mathbf{R}^d \to \mathbf{R}$,

$$f(x) = E(Y \mid X = x), \qquad x \in \mathbf{R}^d,$$

where (Y, X) is distributed as (Y_i, X_i) for any $i = 1, \ldots, n$.

Time series data Our examples of regression function estimation shall involve time series data.

- We can use regression function estimation in time series prediction. Let $Z_1, \ldots, Z_T \in \mathbf{R}$ be a time series and assume that we want to estimate the regression function $f_t : \mathbf{R}^t \to \mathbf{R}$,

  $$f_t(z_1^t) = E(Z_{t+1} \mid Z_1^t = z_1^t),$$

  where $t = 1, \ldots, T-1$, $Z_1^t = (Z_1, \ldots, Z_t)$ and $z_1^t = (z_1, \ldots, z_t)$. Denote

  $$Y_i = Z_{i+1}, \qquad X_i = (Z_1, \ldots, Z_i),$$

  $i = 1, \ldots, n = T - 1$. Now the regression function f_t can be written as

  $$f_i(x) = E(Y_i \mid X_i = x).$$

  Note that it is not reasonable to assume that $(Y_1, X_1), \ldots, (Y_n, X_n)$ would be an i.i.d. sequence.

- The previous setting can be simplified. Let $Z_1, \ldots, Z_T \in \mathbf{R}$ be a time series and assume that we want to estimate the regression function $f_t : \mathbf{R}^k \to \mathbf{R}$,

  $$f_t(z_{t-k+1}^t) = E(Z_{t+1} \mid Z_{t-k+1}^t = z_{t-k+1}^t),$$

  where $t = k, \ldots, T-1$, $Z_l^t = (Z_l, \ldots, Z_t)$ and $z_l^t = (z_l, \ldots, z_t)$ for $1 \le l \le t$. Denote

  $$Y_i = Z_{i+1}, \qquad X_i = (Z_{i-k+1}, \ldots, Z_i),$$

$i = k, \ldots, n = T-1$. When the time series $Z_1, \ldots, Z_T$ is stationary, then the sequence (Y_i, X_i), $i = k, \ldots, n$, consists of identically distributed random variables and the regression function f_t is the same for all time indexes t: we can write

$$f(x) = E(Y_i \,|\, X_i = x).$$

3 Portfolio selection

Let $S_t \in (0, \infty)^d$, $t = 0, 1, 2, \ldots, T$, be a vector time series of asset prices. We denote $S_t = (S_t^1, \ldots, S_t^d)$. We want to choose a portfolio vector $b \in \mathbf{R}^d$ at time T so that the wealth is maximized at time $T+1$. A portfolio vector satisfies $b_i \geq 0$ and $\sum_{i=1}^d b_i = 1$. The value b_i gives the proportion of wealth which is invested in asset S_i at time T. Let W_T be the wealth available at time T. When the portfolio vector is b, then the wealth at time $T+1$ is

$$W_{T+1} = W_T \cdot b^T (S_{T+1}/S_T).$$

We want to choose b so that

$$E\left(Y \,|\, X_T = x\right), \qquad Y = \log_e(b^T(S_{T+1}/S_T))$$

is maximized, where X_T is the relevant information available at time T. Next we construct the regression data. Let

$$U_t = S_t/S_{t-1} = (S_t^1/S_{t-1}^1, \ldots, S_t^d/S_{t-1}^d), \qquad t = 1, 2, \ldots,$$

be the price relatives. We assume that the U_t are a realization of an identically distributed process. Let

$$Y_t = \log_e(b^T U_{t+1}) \in \mathbf{R}, \qquad t = 0, \ldots, T-1,$$

be the realizations of the response variable. The explanatory variables can be chosen in many ways. We shall use the past observations to choose the optimal portfolio. Let $k \geq 1$ be an integer, and let

$$X_t = (U_{t-k+1}, \ldots, U_t) \in \mathbf{R}^{dk}, \qquad t = k, \ldots, T.$$

Our regression data is

$$(X_t, Y_t), \qquad t = k, \ldots, T-1.$$

We assume that the regression data is identically distributed and denote with (X, Y) a random variable distributed as (X_t, Y_t). Note that Y depends on b. We estimate the regression function

$$f(x; b) = E(Y \mid X = x), \qquad x \in \mathbf{R}^{dk}.$$

We choose the optimal portfolio vector $\hat{b}$ as

$$\hat{b} = \mathrm{argmax}_{b \in S_d} \hat{f}(X_T; b),$$

where

$$S_d = \left\{ b \in \mathbf{R}^d : b_i \geq 0, \ \sum_{i=1}^{d} b_i = 1 \right\}.$$

4 Illustration

We look at the following code in

http://cc.oulu.fi/ jklemela/finatool/

```
# First load the library

library(finatool)

# we use package "Rdonlp2" for the optimization

library(Rdonlp2)

Portfolio of stocks

# first we download data

ticker<-c("^GDAXI","^MDAXI")
destfile<-"~/pois"
ry<-read.yahoo(ticker, source="web", destfile=destfile)

destfile<-"/Users/jsk/pois"
ry<-read.yahoo(ticker, source="web", destfile=destfile)

dm<-data.manip(ry,ticker)
plot(dm$dat[,1],type="l")
```

```
x11()
plot(dm$dat[,2],type="l")

method<-rep("relative",length(ticker))
df<-data.final(dm,ticker,method=method)
plot(df)

# we shall use the previous price relatives to predict future price relatives

d<-length(ticker)
n<-dim(df)[1]
U<-matrix(0,n,d)
for (i in 1:d) U[,i]<-df[1:n,i]

k<-4
mp<-make.portdat(U,k)

# we apply the nearest neigborhood method with m nearest neighbors

m<-100
wts<-pf.nn(mp$Z,mp$arg,mp$X,m=m)

# we apply the kernel method with smoothing parameter h

h<-1
wts<-pf.kernel(mp$Z,mp$arg,mp$X,h=h)

# we study the historical performance of the nearest neighborhood method

estimator<-"nn"
m<-15
pfseq<-pf.seq(mp$Z,mp$X,estimator=estimator,m=m)

end<-length(pfseq$wealth)
start<-1 #end-round(n/2)
plot(pfseq$wealth[start:end]/pfseq$wealth[start])

plot(pfseq$port[start:end,1])

plot(pfseq$return[start:end])
```

```
# we compare the nn-portfolio choice to the equally weighted portfolio

method<-rep("price",length(ticker))
dp<-data.final(dm,ticker,method=method)
dpmean<-(dp[,1]+dp[,2])/2

mata<-matrix(0,end-start+1,2)
mata[,1]<-dpmean[start:end]/dpmean[start]
mata[,2]<-pfseq$wealth[start:end]/pfseq$wealth[start]
matplot(mata,type="l",xlab="time",ylab="wealth")


# we study the historical performance of the kernel method

estimator<-"kernel"
h<-0.1
kernel<-"gauss"
pfseq.ker<-pf.seq(mp$Z,mp$X,estimator=estimator,h=h,kernel=kernel)

end<-length(pfseq$wealth)
start<-end-round(n/2)
plot(pfseq$wealth[start:end]/pfseq$wealth[start])

# we compare the kernel-portfolio to the nn-portfolio choice and
# to the equally weighted portfolio

mata<-matrix(0,end-start+1,3)
mata[,1]<-dpmean[start:end]/dpmean[start]
mata[,2]<-pfseq$wealth[start:end]/pfseq$wealth[start]
mata[,3]<-pfseq.ker$wealth[start:end]/pfseq.ker$wealth[start]
matplot(mata,type="l",xlab="time",ylab="wealth")
```

5 Examination

A possible question in the examination:

1) Explain how regression function estimation can be used to choose a portfolio.

A possible practical work:

- Study portfolio choice making similar computations as in Section 4. Choose suitable assets (stocks, indexes, currencies,...) and choose a regression function estimator (kernel estimator, nearest neighborhood estimator,...). Study the impact of the parameter k and the parameters of the estimator to the portfolio choice.