Lecture 4

Jussi Klemelä

February 10, 2009

1 Single index models

There exists a large number of regression function estimators which are based on the single index model assumption

$$f(x) = g(x^T\theta), \qquad x \in \mathbf{R}^d, \tag{1}$$

where $g : \mathbf{R} \to \mathbf{R}$ is the unknown link function and $\theta \in \mathbf{R}^d$ is the unknown index vector. Note that unlike in the generalized linear model, to be defined later, the link function g is now unknown and needs to be estimated. An index summarizes many variables into one number, for example stock index, inflation index, cost-of-living index, or price index are such indeces. A linear index $x^T\theta$ can be generalized to a non-linear index $v_\theta(x)$, depending on parameter θ, and then the single index model can be written as

$$f(x) = g\left(v_\theta(x)\right), \qquad x \in \mathbf{R}^d.$$

We shall consider linear single index models (1).

Identification. The vector θ is not uniquely defined: the use of the vector $c\theta$ and the link function $g_c(u) = g(u/c)$, with some $c > 0$, leads to the same regression function f. To assure the uniqueness we shall assume that $\|\theta\| = 1$. (We could also assume that the first component of θ is equal to one.) Also, the sign of the coefficient vector θ is not unique, because the use of the vector $-\theta$ and the link function $g_-(u) = g(-u)$ leads to the same regression function.

Estimation. For a given θ, the link function g can be estimated by applying univariate nonparametric regression with $X_i^T\hat{\theta}$, $i = 1, \ldots, n$, as observations of a univariate explanatory variable. Thus we can proceed in the estimation of the regression function in the single-index model by first estimating the parameter vector θ and then estimating the link function g. We consider both the minimization estimation approach and the average derivative approach.

M-estimation approach. In the minimization-estimation (M-estimation) approach one finds for each fixed θ a nonparametric estimator $\hat{g}_\theta$ of $g_\theta(t) = E(Y_1 \mid X_1^T\theta = t)$ and then estimates θ by

$$\hat{\theta} = \text{argmin}_\theta \sum_{i=1}^{n} \psi\left(Y_i, \hat{g}_\theta(X_i^T\theta)\right),$$

where ψ is a contrast function. The contrast function ψ can be chosen for example as

$$\psi(y, z) = |y - z|^2,$$

which leads to a semiparametric least squares estimator.

2 Nonparametric portfolio selection

2.1 The approach

We have available d assets and we want to choose an optimal allocation of the wealth among these assets. Let us denote with T the current time. We want to choose the portfolio so that the wealth is maximized at time $T + 1$. We have available the previous asset prices $S_t \in (0, \infty)^d$, $t = 0, 1, \ldots, T$, which can be used in portfolio selection. We denote $S_t = (S_t^1, \ldots, S_t^d)$.

A portfolio vector $b \in \mathbf{R}^d$ 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 shall choose b so that the regression function

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

is maximized, where X_T is the relevant information available at time T and $u : (0, \infty) \to \mathbf{R}$ is a utility function. We discuss various utility functions in Section 2.2.

Response variables. Let

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

be the price relatives. We assume that the U_t are identically distributed random variables. Let

$$Y_t = u\left(b^T U_{t+1}\right) \in \mathbf{R}, \qquad t = 0, \dots, T-1,$$

be the realizations of the response variable.

Explanatory variables. We denote the explanatory data by

$$X_t \in \mathbf{R}^p, \qquad t = t_0, \dots, T,$$

and this data can consist of technical or fundamental information. For example, we can use the past observations to choose the optimal portfolio. Then

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

where $k \geq 1$ is an integer.

Portfolio selection. Our regression data is

$$(X_t, Y_t), \qquad t = t_0, \dots, T-1.$$

We can 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

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

with the estimator $\hat{f}(x; b)$. We choose the optimal portfolio vector $\hat{b}$ as

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

where

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

2.2 Utility functions

The portfolio selection problem can be stated in the following way. At time t we have available wealth $W_t > 0$ and a collection B of trading strategies. Our investment horizon extends to time T in the future. To each trading strategy $b \in B$ we associate final wealth $W_T(b)$ at time T. Wealth $W_T(b)$ is random variable, whose distribution is unknown. We have to use statistical methods to get a hold on the distribution of $W_T(b)$. For the moment, let us discard the fact that the distribution of $W_T(b)$ is unknown. Thus we have a collection of random variables

$$U(b) = \frac{W_T(b)}{W_t}, \qquad b \in B,$$

and we choose such trading strategy b that the distribution of $U(b)$ is most advantageous.

For example, in the one step long only portfolio selection the strategies b are the portfolio weights: $b = (b_1, \ldots, b_d)$, $b_i \geq 0$, $\sum_{i=1}^{d} b_i = 1$, where d is the number of available assets. Now

$$U(b) = b^T U_T, \tag{2}$$

$U_T = (S_T^1/S_t^1, \ldots, S_T^d/S_t^d)$, where $S^1, \ldots, S^d$ are the d available assets.

Comparison of distributions. The selection of the best trading strategy requires that we can make an ordering among the univariate random variables $U(b)$ (or an ordering among their distributions). Figure 1 illustrates the comparison of distributions. Frame a) shows two densities of the return distribution which are easy to compare; the densities have the same shape but the other dominates the other, because its mode is at 1.2 wheras the mode of the other density is at 1.05. Frame b) shows two densities which are not straightforward to compare; the mode of the other is at 1.2 but its variance is larger, whereas the mode of the other is at 1.05 but its variane is smaller.

Utility functions. A useful way to order distributions is to order they according to the value of expected utility:

$$Eu(U(b)),$$

where $u : (0, \infty) \to \mathbf{R}$ is a utility function. It is natural to assume that a utility function is strictly increasing and strictly concave. A utility function is increasing because one prefers a larger wealth to a lesser wealth. A utility

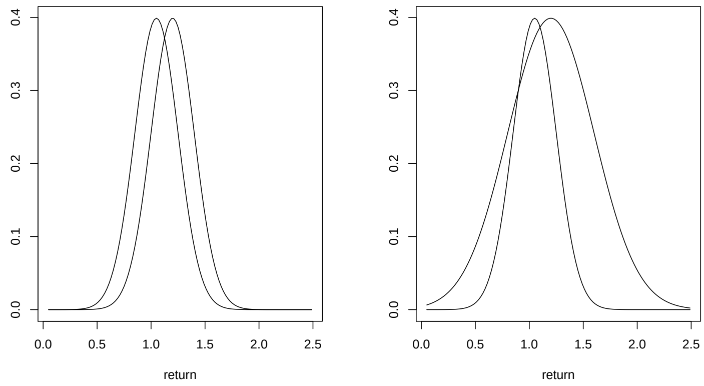

Figure 1: *(Comparison of distributions.)* Frame a) shows two return densities which are easy to compare and frame b) shows two return densities which are difficult to compare.

function is concave because when the wealth increases, the value of additional wealth declines. The utility function corresponding to the constant relative risk aversion (CRRA) preferences is

$$u(w) = \begin{cases} \frac{w^{1-\gamma}}{1-\gamma}\,, & \text{if } \gamma > 1, \\ \log_e w\,, & \text{if } \gamma = 1. \end{cases} \tag{3}$$

An other family of utility functions is

$$u(w) = 1 - e^{-\alpha w}, \qquad \alpha > 0. \tag{4}$$

Figure 2 shows CRRA utility functions and exponential utility functions (4) for different parameter values. We show the normalized utility functions

$$\tilde{u}(w) = \frac{u(w) - u(1)}{u(2) - u(1)}$$

so that $\tilde{u}(1) = 0$ and $\tilde{u}(2) = 1$. Note that the ordering of the distributions is not affected by linear transformations $au(w) + c$, $a > 0$, $c \in \mathbf{R}$, because

$$E[au(U(b)) + c] = aEu(U(b)) + c.$$

The figure shows that larger values of γ or α are used when one is more risk averse, because the curvature of the utility functions increases when γ or α are increased.

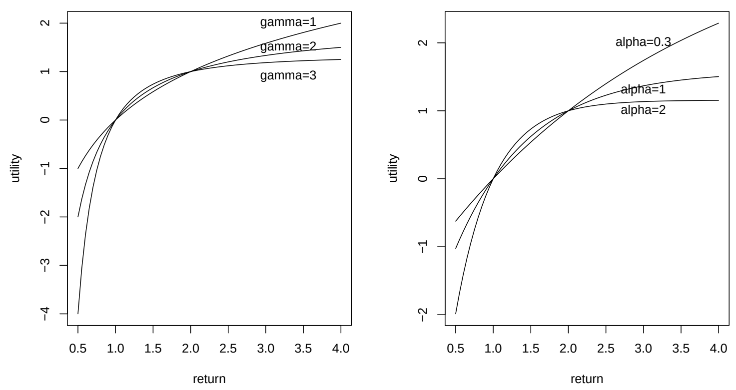

Figure 2: *(Utility functions.)* Frame a) shows CRRA utility functions for parameter values $\gamma = 1$, $\gamma = 2$, and $\gamma = 3$. Frame b) shows utility functions in (4) for parameter values $\alpha = 0.3$, $\alpha = 1$, and $\alpha = 2$.

Figure 3 illustrates the ranking of distributions according to expected utility, when the utility function is the CRRA function with $\gamma = 1$. The densities are Gaussian with $\mu = 1.05$, $\sigma = 0.05$ (red density), $\mu = 1.1$, $\sigma = 0.05$ (blue density), $\mu = 1.05$, $\sigma = 0.2$ (black density). The black density is the worst and the blue density is the best, as is clearly reasonable.

Markowitz utility. Portfolio choice with mean-variance preferences ranks the distributions according to

$$EU(b) - \frac{\gamma}{2} \mathrm{Var}[U(b)],$$

where $\gamma \geq 0$ is the coefficient of absolute risk aversion. Parameter γ measures the investor's absolute risk aversion $u''(w)/u'(w)$.

When we consider the one step long only portfolio choice, then $U(b)$ is defined in (2) and we can solve the portfolio weights explicitely:

$$b = \Sigma^{-1} 1_d \, \frac{\gamma W_t - 1_d \cdot \Sigma^{-1}\mu}{\gamma W_t 1_d \cdot \Sigma^{-1} 1_d} + \frac{\Sigma^{-1}\mu}{\gamma W_t},$$

where $\mu = EU_T$, $\Sigma = \mathrm{Var}[U_T]$, and $1_d = (1, \ldots, 1) \in \mathbf{R}^d$.

3 Illustrations

We look at the following code in

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

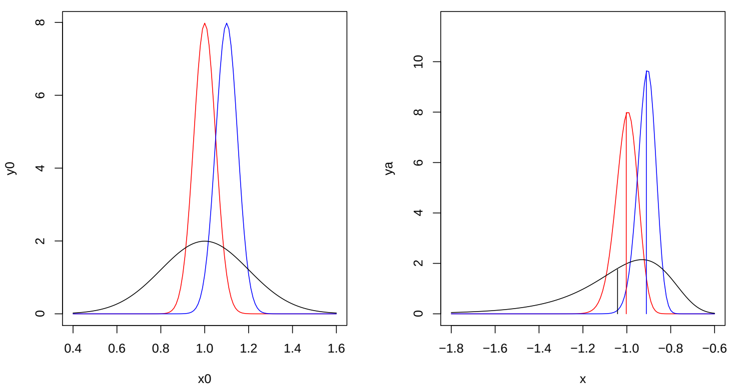

Figure 3: *(Comparison of distributions.)* Frame a) shows three density functions of return U and frame b) shows the density functions of $u(U)$, when the utility function u is the CRRA function with $\gamma = 1$. The expectetations of $u(U)$ are marked with vertical vectors. The densities are Gaussian with $\mu = 1.05$, $\sigma = 0.05$ (red density), $\mu = 1.1$, $\sigma = 0.05$ (blue density), $\mu = 1.05$, $\sigma = 0.2$ (black density).

```
# single index regression

# we use package "Rdonlp2" for the optimization

library(Rdonlp2)

# we obtain returns of the DAX stock index

ticker<-c("^GDAXI")
destfile<-"~/pois"
# destfile<-"C:\\Documents and Settings\\user\\Desktop\\pois"
ry<-read.yahoo(ticker, source="web", destfile=destfile)
#save(file="/home/jsk/Arti/statfina/Dax.var",list=c("ry"))
#load(file="/Users/jsk/Karhu/Arti/statfina/DaxMdax.var")
dm<-data.manip(ry,ticker)
method<-"return"
df<-data.final(dm,ticker,method=method)
n<-dim(df)[1]
S<-matrix(df[1:n,1],n,1)
plot(S,type="l")
```

```
# we calculate volatilities for the 5 day periods

perlen<-5
pernum<-floor(n/perlen)
volas<-matrix(0,pernum,1)
for (i in 1:pernum){
    beg<-(i-1)*perlen+1
    end<-(i-1)*perlen+perlen
    period<-S[beg:end]
    volas[i]<-sqrt(sum(period^2)/perlen)*sqrt(252)
}
plot(volas,type="l")

# we use now the volatilities of d periods to predict
# the volatility of the next period

d<-2
dendat<-matrix(0,pernum-d,d+1)
for (i in 1:(pernum-d))
    for (j in 1:(d+1))
      dendat[i,j]<-volas[i+j-1]

# we make the copula transform for the explanatory variables
# and estimate the regression function

x<-copula.trans(dendat[,1:d])
plot(x)
y<-dendat[,d+1]

# we calculate the best index

h<-0.6
thetahat<-single.index(x,y,h=h,kernel="gauss")

# the final estimate is a kernel estimate

z<-x%*%thetahat
h<-0.4
N<-60
pcf<-pcf.kernesti(z,y,h,N)
```

```
dp<-draw.pcf(pcf)          # requires package "denpro"
matplot(dp$x,dp$y,type="l",xlim=c(-3,3),ylim=c(0,1.1))
matplot(z,y,add=TRUE)


# portfolio selection

# first we download data

ticker<-c("^GDAXI","^MDAXI")
destfile<-"~/pois"
# destfile<-"C:\\Documents and Settings\\user\\Desktop\\pois"
ry<-read.yahoo(ticker, source="web", destfile=destfile)

#save(file="/home/jsk/Arti/statfina/var/DaxMdax.var",list=c("ry"))
#save(file="/Users/jsk/Karhu/Arti/statfina/var-ada/DaxMdax.var",list=c("ry"))
#load(file="/Users/jsk/Karhu/Arti/statfina/var-ada/DaxMdax.var")

dm<-data.manip(ry,ticker)

method<-rep("price",length(ticker))
dfs<-data.final(dm,ticker,method=method)
plot(dfs[,1],type="l")

plot(dfs[,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
marginal<-"gauss"
mp<-make.portdat(U,k,marginal=marginal,rate=0.02/360)
```

```
# we study the historical performance of the nearest neighborhood method

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

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

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")
```

4 Examination

A possible question in the examination:

3) a) Define the single index model.
   b) Explain how the regression function can be estimated in the single index model using M-estimation.