CATÓLICA-LISBON Forecasting Lab | NECEP

NECEP Research Papers

Robust Filtering with Quantile Regression

João Borges de Assunção

Pedro Afonso Fernandes

Research Paper 2/2022

Robust Filtering with Quantile Regression

João Borges de Assunção
Pedro Afonso Fernandes

Universidade Católica Portuguesa
Católica Lisbon School of Business & Economics
Católica Lisbon Research Unit of Business and Economics
Católica Lisbon Forescasting Lab (NECEP)

February 21, 2022

Abstract

This working paper proposes a new, practical method to compute the non-linear Mosheiov-Raveh (MR) filter using least absolute deviations (LAD) instead of the linear programming approach proposed by these two authors. This paper is embodied with an implementation in the R programming language of the proposed method which facilitates the computation of the MR filter in current applications to produce a robust estimate, namely, of the GDP trend growth. This technique may be appropriate to deal with non linear time series or structural changes.
Keywords: Business cycles, Non linear time series, Robust filtering, Software.

1 Introduction

In Ordinary Least Squares (OLS), the fitted values $\hat{y}_i = X_i\hat{\beta}$ are the *conditional mean* of the dependent variable on the values of the independent variables. In **median regression**, the fitted values are the *conditional median* of the dependent variable. Thus, median regression choose $\hat{\beta}$ to minimize the sum of absolute residuals:

$$\min \sum_{i=1}^{N} |y_i - x_i\beta| \tag{1}$$

This problem, known as **Least Absolute Deviations (LAD)**, can be solved with the simplex and other methods, namely, interior point methods from the following linear programming representation (Portnoy and Koenker, 1997):

$$\begin{gathered} \min \left\{ 1^T u + 1^T v \right\} \\ \text{s.t. } y = X\beta + u - v \\ (u, v) \in \mathbb{R}_+^{2N} \end{gathered} \tag{2}$$

where 1 denotes a column vector of ones with N elements.

More generally, the **quantile regression** predicts the conditional τ-quantile of the dependent variable, where the median regression corresponds to $\tau = 0.5$. Several packages are available to compute this regression, namely, the R package **quantreg** (Koenker, 2019), function rq(y ~ x, tau = 0.5, data).

In this paper, we propose a new method to compute the Mosheiov-Raveh (1997) filter using LAD and quantile regression. This method is easy to use and more flexible than the traditional linear programming approach to compute the MR filter. It also relaxes the monotonicity condition on the piecewise linear trend.

2 The Mosheiov-Raveh filter

For a time series $\{y_t\}$, $t = 1, \ldots N$, the Mosheiov-Raveh (1997) piecewise linear trend $h \equiv \{h_t\}$ is chosen either to minimize the sum of the absolute deviations from the data, a possible measure of fidelity, or the sum of its second differences, a measure of its smoothness also in the ℓ_1 norm:

$$\min \left\{ \sum_{t=1}^{N} |y_t - h_t| + \theta \sum_{t=1}^{N-2} |(h_{t+2} - h_{t+1}) - (h_{t+1} - h_t)| \right\} \tag{3}$$

where the penalty θ might be fixed as the squared root of the Hodrick-Prescott (1997) parameter λ, namely, $\theta = \sqrt{1600} = 40$ for quarterly data (Ravn and Uhlig, 2002). It is convenient to express the MR problem in matrix form

$$\min \{ \|y - I_n h\|_1 + \theta \|Dh\|_1 \} \tag{4}$$

where $\|u\|_1 = \sum_i |u_i|$ denotes the ℓ_1 norm of vector u and $D \in \mathbb{R}^{(N-2)\times N}$ is an upper triangular Toeplitz matrix with first row $[1 \; -2 \; 1 \; 0 \; \cdots \; 0]$ (kim et al., 2009). Like LAD, this problem has a linear programming representation (Mosheiov and Raveh, 1997):

$$\begin{aligned} &\min \left\{ 1^T u + 1^T v + \theta 1^T a + \theta 1^T b \right\} \\ &\text{s.t. } y = I_N h + u - v \\ &\quad Dh - a + b = 0 \\ &\quad h - L(h) \geq 0 \\ &\quad (u, v) \in \mathbb{R}_+^{2N} \\ &\quad (a, b) \in \mathbb{R}_+^{2(N-2)} \end{aligned} \tag{5}$$

where L is the lag operator, that is, $L(h_t) = h_{t-1}$. Solving this problem with the simplex can be arduous, namely, with thousands of observations. Additionally, we need to impose a condition on the monotonicity of the trend, namely, $h_t \geq h_{t-1}$ or $h_t \leq h_{t-1}$, for $t = 2, \ldots, N$, which may be irrealistic for common applications in macroeconomics.

3 Computing the MR filter with quantile regression

Relaxing the previous monotonicity condition, the linear programme of the MR filter can be expressed as

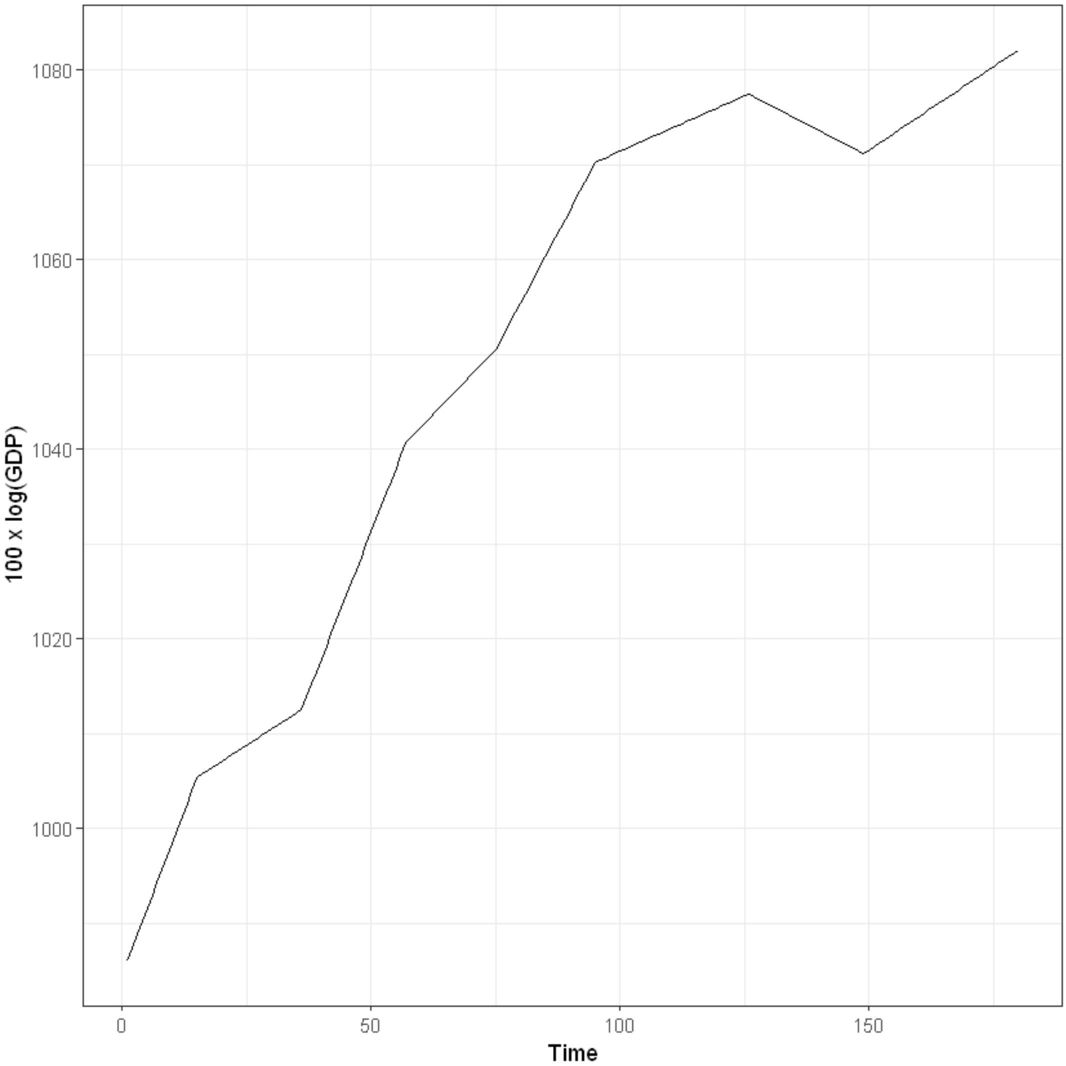
$$\begin{aligned} &\min \left\{ 1^T \tilde{u} + 1^T \tilde{v} \right\} \\ &\text{s.t. } \tilde{y} = Xh + \tilde{u} - \tilde{v} \\ &\quad (\tilde{u}, \tilde{v}) \in \mathbb{R}_+^{2(2N-2)} \end{aligned} \tag{6}$$

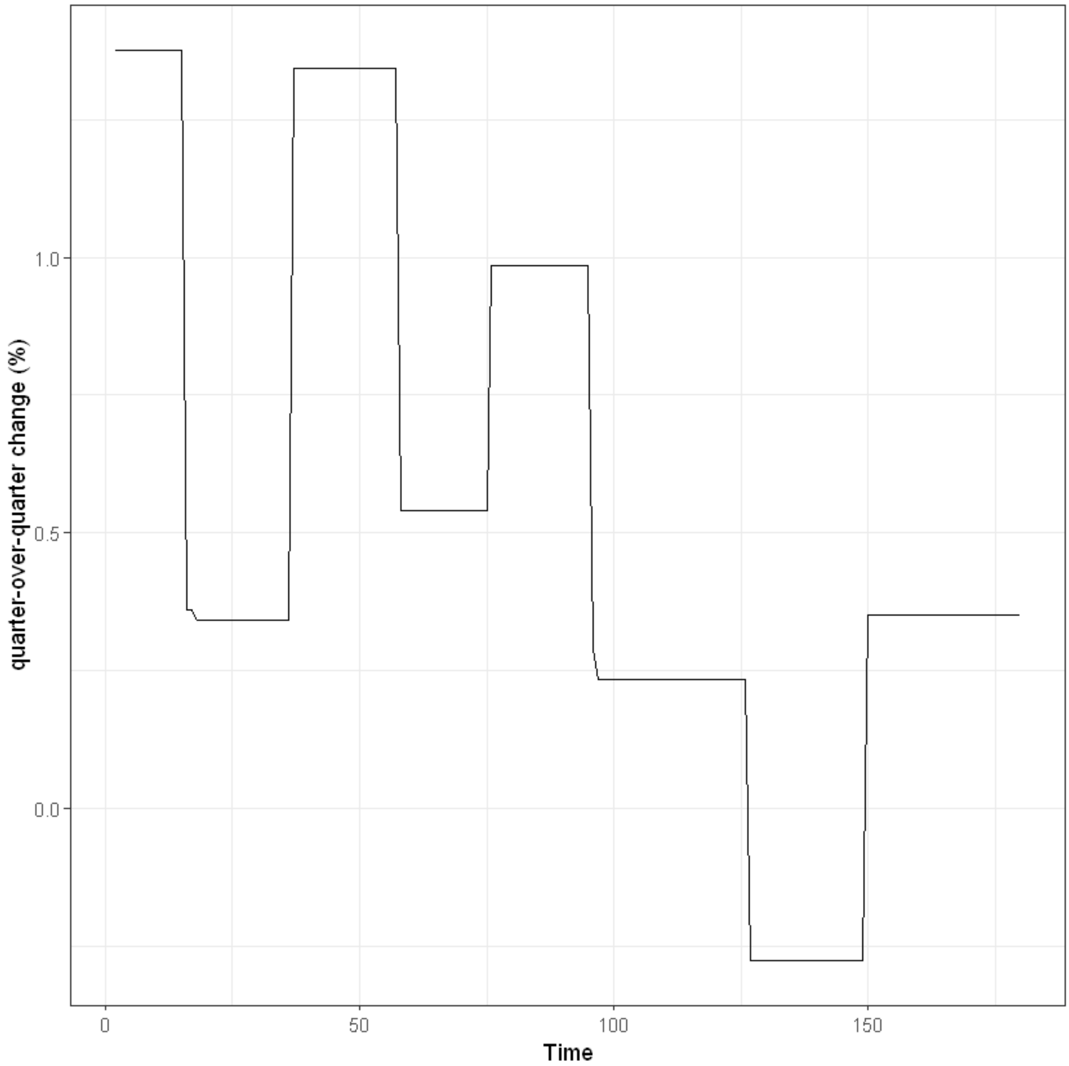
where $\tilde{u}^T = [u|b]$, $\tilde{v}^T = [v|a]$, $X^T = [I_N|\theta D]$ and $\tilde{y}^T = [y|0]$ with 0 as a vector of $N-2$ zeros. Given this reformulation of the problem, the MR trend h corresponds simply to the estimated N-coefficients of the median regression of $\tilde{y}$ on X that solves the problem:

$$\min \sum_{i=1}^{2N-2} |\tilde{y}_i - x_i h| \tag{7}$$

The following application in R computes the MR filter for the Portuguese 100 times the log of GDP from 1977Q1 to 2021Q4 using this approach:

```r
library(MASS)
library(quantreg)
library(ggplot2)

theta <- 40

data <- read.csv("gdp.csv")
n <- nrow(data)
y <- 100*log(data$GDP_PT)

time <- 1:n

V <- rep(0, n)
V[1] <- 1
V[2] <- -2
V[3] <- 1

T <- toeplitz(V)
T[lower.tri(T)] <- 0
m <- n-2
D <- T[1:m,]

X <- rbind(diag(n), theta * D)
y_tilde <- c(y, rep(0,m))

dt <- data.frame(y_tilde,X)

model <- c("y_tilde ~ 0")
for (i in 1:n) model <- paste(model, "+X", i, sep = "")

mod1 <- rq(model, tau = 0.5, data = dt)
mrt1 <- coef(mod1)
dmrt1 <- mrt1[2:n] - mrt1[1:(n-1)]

write.csv(mrt1,"mr_out1.csv")

qplot(y = mrt1, x = time, geom = "path") +
labs(x = "Time", y = "100 x log(GDP)" ) + theme_bw()

qplot(y = dmrt1, x = time[2:n], geom = "path") +
labs(x = "Time", y = "quarter-over-quarter change (%)" ) + theme_bw()
```

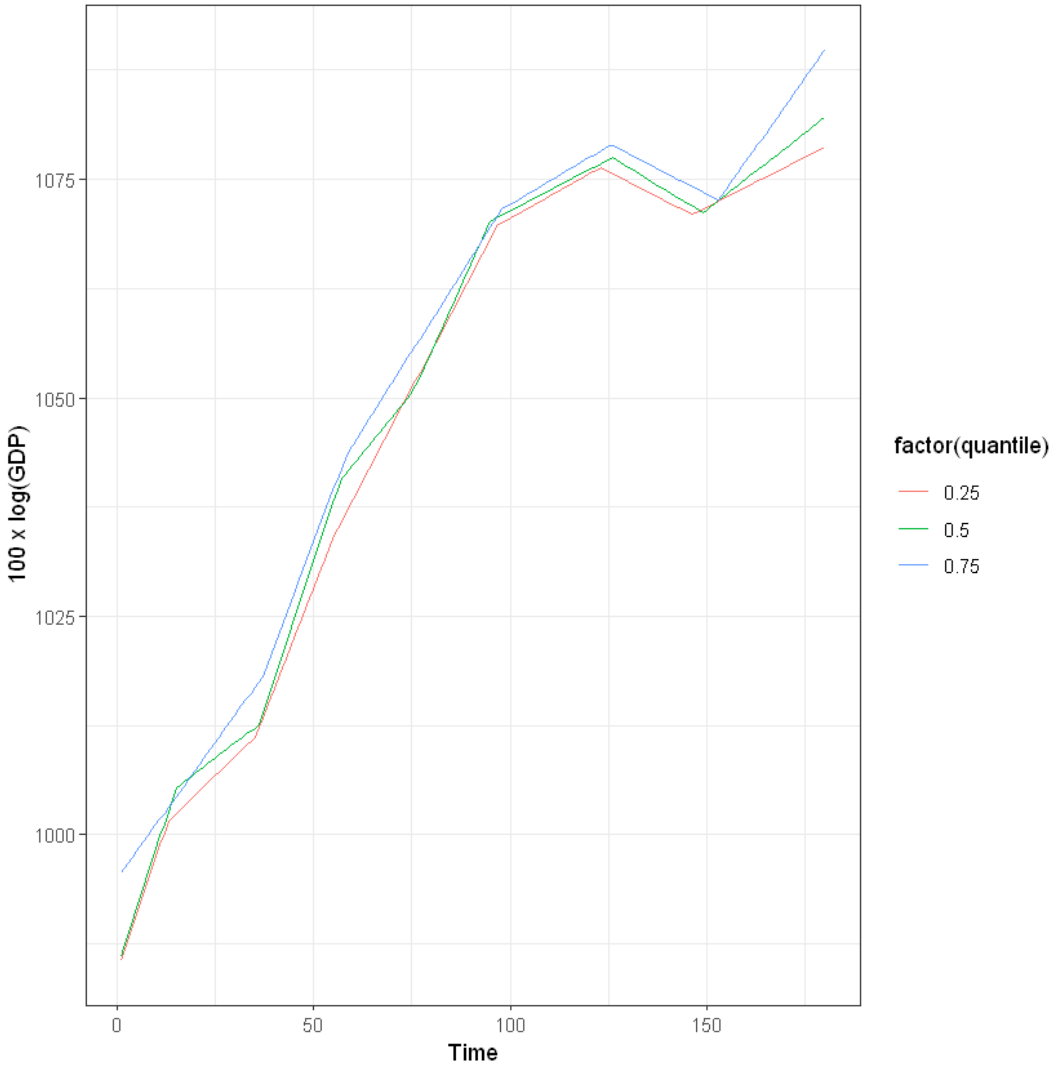

1080
1060
1040
1020
1000
100 x log(GDP)
0
50
100
150
Time

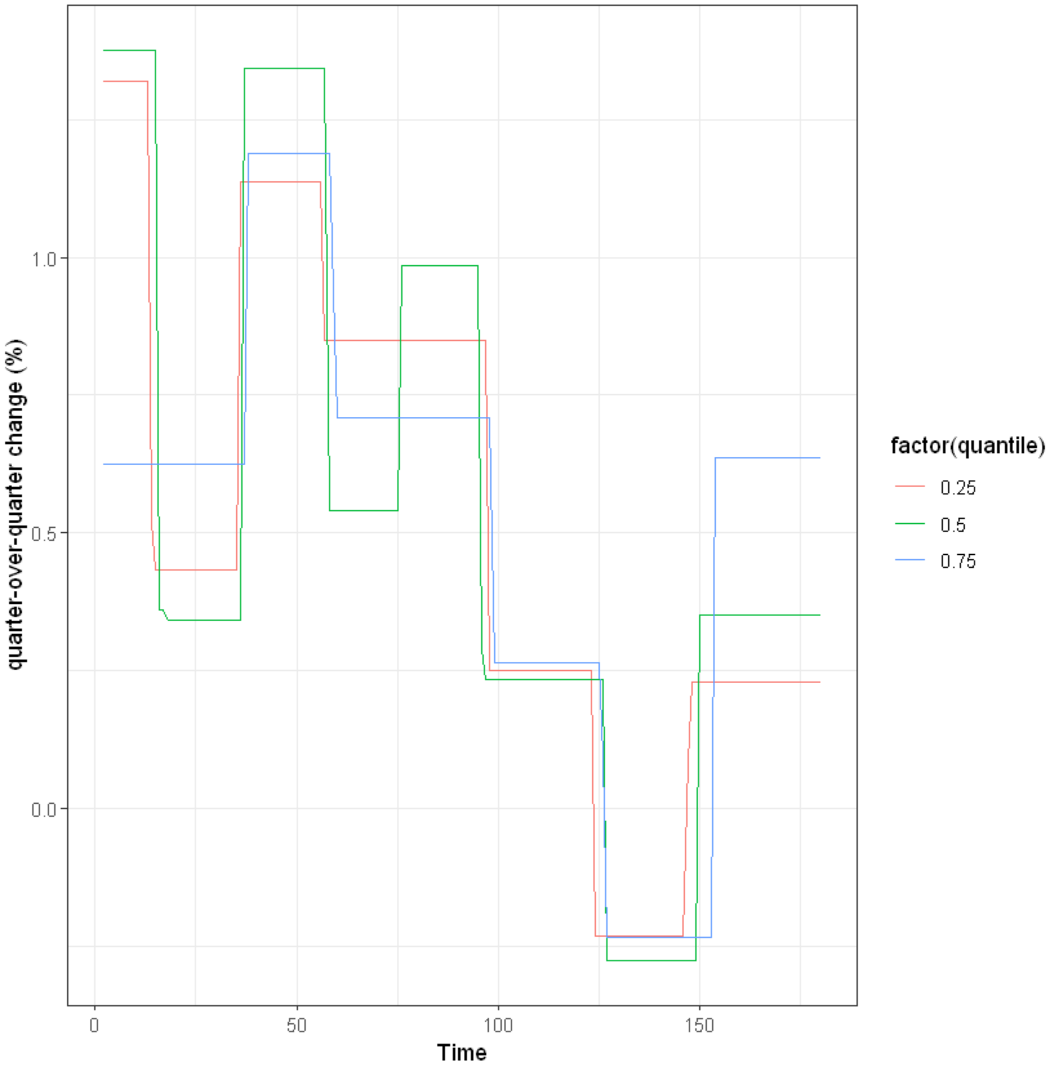


One interesting feature of this approach is the possibility to compute the MR filter not only for the median ($\tau = 0.5$) but also for other quantiles, namely, for the first and third quartiles (respectively, $\tau = 0.25$ and $\tau = 0.75$).

```
mod2 <- rq(model, tau = 0.25, data = dt)
mrt2 <- coef(mod2)
dmrt2 <- mrt2[2:n] - mrt2[1:(n-1)]

mod3 <- rq(model, tau = 0.75, data = dt)
mrt3 <- coef(mod3)
dmrt3 <- mrt3[2:n] - mrt3[1:(n-1)]

mrt <- c(mrt2,mrt1,mrt3)
quantile <- c(rep(0.25,n),rep(0.50,n),rep(0.75,n))
time2 <- c(time,time,time)
qplot(y = mrt, x = time2, geom = "path", col = factor(quantile)) +
labs(x = "Time", y = "100 x log(GDP)") + theme_bw()
```

```
dmrt <- c(dmrt2,dmrt1,dmrt3)
quantile <- c(rep(0.25,(n-1)),rep(0.50,(n-1)),rep(0.75,(n-1)))
time3 <- c(time[2:n],time[2:n],time[2:n])
qplot(y = dmrt, x = time3, geom = "path", col = factor(quantile)) +
labs(x = "Time", y = "quarter-over-quarter change (%)") + theme_bw()
```





4 Duality

As the median or quantile regression for $\tau = 0.5$, the MR filter has the following dual formulation (Portnoy and Koenker, 1997):

$$\begin{aligned} &\max \left\{ \tilde{y}^T d \right\} \\ &\text{s.t. } X^T d = 0 \\ &\quad d \in [-0.5, 0.5]^{2N-2} \end{aligned} \tag{8}$$

or, setting $a = d + 0.5$

$$\begin{aligned} &\max \left\{ \tilde{y}^T a \right\} \\ &\text{s.t. } X^T a = 0.5 X^T 1 \\ &\quad a \in [0, 1]^{2N-2} \end{aligned} \tag{9}$$

Adding slack variables s satisfying the constraint $a + s = 1$, we obtain the logarithmic barrier potential-function:

$$B(a, s, \mu) = \tilde{y}^T a + \mu \sum_{i=1}^{2N-2} (\log a_i + \log s_i) \tag{10}$$

which should be maximized subject to the constraints $X^T a = 0.5X^T 1$ and $a + s = 1$. Applying the Newton's method to this problem and noting that $\mu \to 0$ in the optimum, Portnoy and Koenker (1997) found the following analytical solution for the primal variable h which is the MR trend in our case:

$$h = (X^T W X)^{-1} X^T W \tilde{y} \tag{11}$$

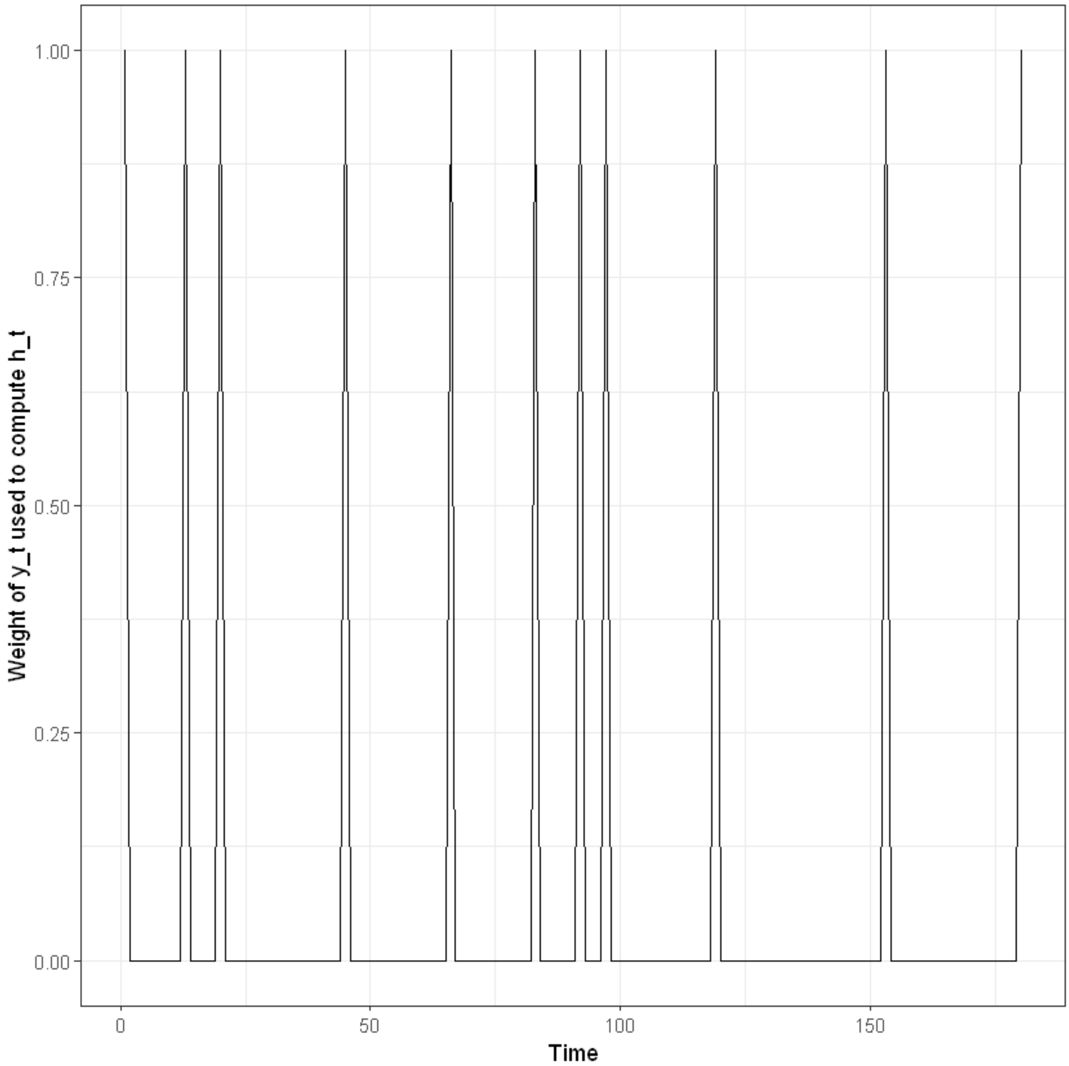
where $W = (A^{-2} + S^{-2})^{-1}$ with A and S being diagonal matrices whose main diagonals are the values of the dual variables a and s, respetively, and zero otherwise.

The following application deals with the zero values in some elements of these diagonals by introducing a small (tiny) value:

```
tv <- 0.000001
a  <- mod1$dual
mn <- length(a)
a  <- pmax(a,rep(tv,mn))
a  <- pmin(a,rep(1-tv,mn))
A2 <- diag(a^(-2))
s  <- rep(1,mn) - a
S2 <- diag(s^(-2))
W  <- solve(A2+S2)
M  <- solve(t(X) %*% W %*% X) %*% t(X) %*% W
M2 <- M[,1:n]
h  <- M %*% y_tilde

write.csv(M,"M.csv")
write.csv(M2,"M2.csv")
write.csv(h,"mr_out2.csv")
```

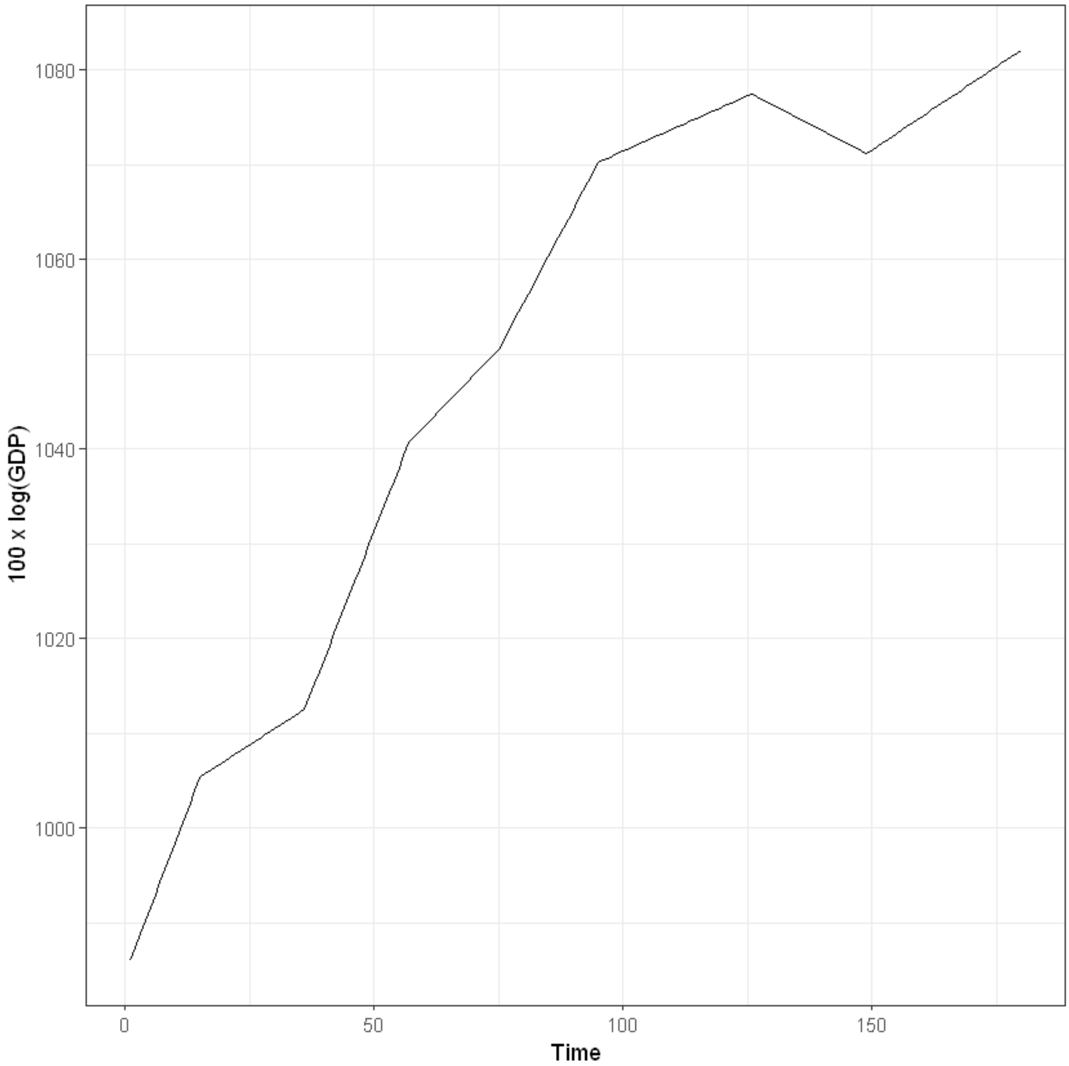
Each row i of the previous matrix M2, which corresponds to the N first columns of the $N \times (2N-2)$ matrix $M = (X^T W X)^{-1} X^T W$, contains the weigths associated with each observation y_t used to compute the MR trend level in period i with $i, t = 1, \ldots, N$. The following figure represents the main diagonal elements of this matrix M2 and suggests that the MR filter is constructed with a small set of critical observations such that $h_t = y_t$:

```
qplot(y = diag(M2), x = time, geom = "path") +
labs(x = "Time", y = "Weight of y_t used to compute h_t" ) + theme_bw()
```



The following figure represents the MR trend h_t computed with this dual approach which is identical with the trend already computed with quantile regression, as described above:

```
qplot(y = h, x = time, geom = "path") +
labs(x = "Time", y = "100 x log(GDP)" ) + theme_bw()
```



5 Computing the MR filter with the Meketon's algorithm

The LAD problem can be solved with interior linear programming techniques, namely, with the simple affine scalling algorithm proposed by Meketon (1986). This algorithm is easy to code, especially if a weighted least squares subroutine is available such as the function lm.wfit(x,y,w) from the R package **stats** (loaded by default). In fact, it can be viewed as an iterative reweighted least squares algorithm. The following implementation is available in (Koenker, 2008):

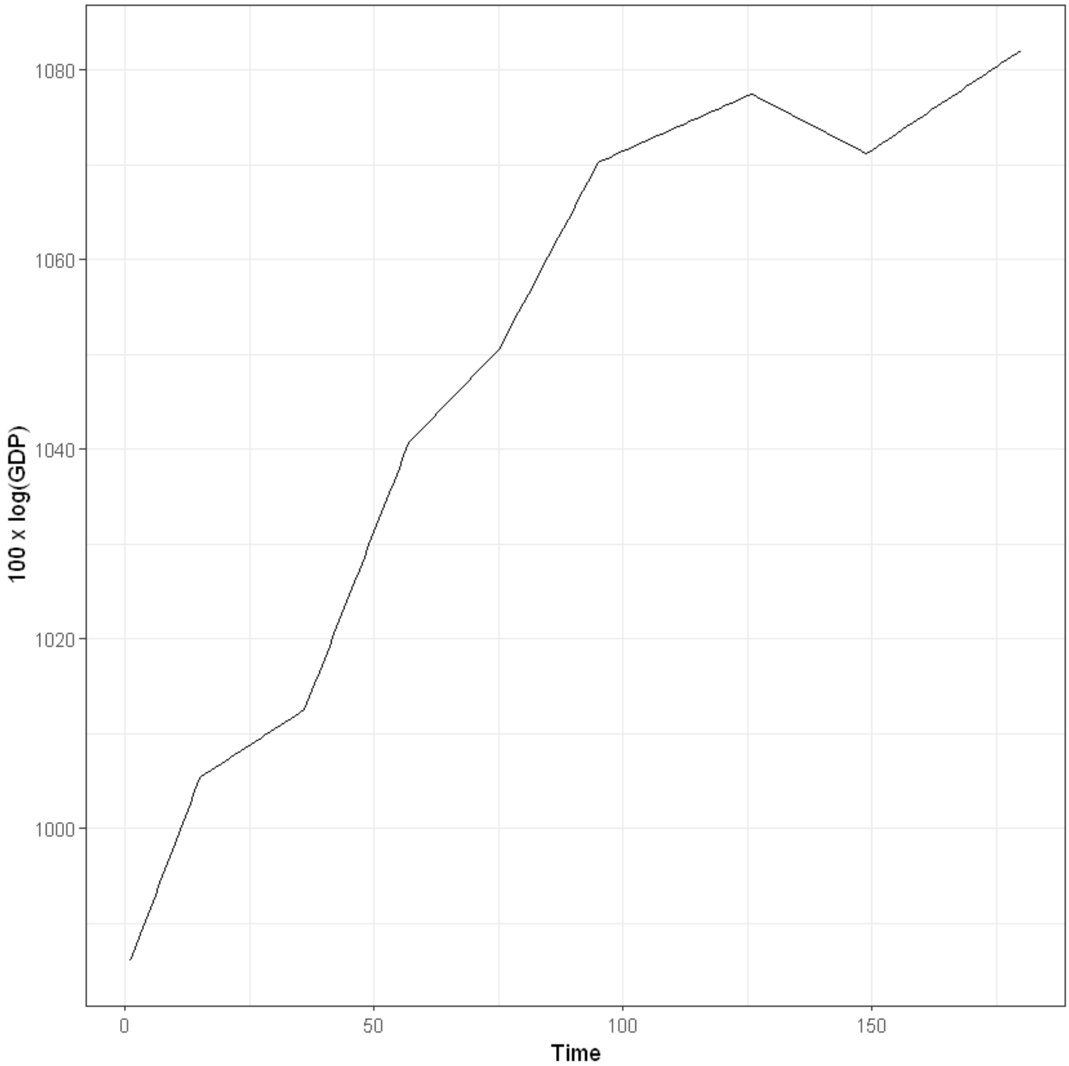
```
meketon <- function (y, x, eps = 1e-04, beta = 0.97)
{
    f <- lm.fit(x,y)
    n <- length(y)
    w <- rep(0, n)
    d <- rep(1, n)
    its <- 0
    while(sum(abs(f$resid)) - crossprod(y, w) > eps)
```

```
    {
        its <- its + 1
        s <- f$resid * d
        alpha <- max(pmax(s/(1 - w), -s/(1 + w)))
        w <- w + (beta/alpha) * s
        d <- pmin(1 - w, 1 + w)^2
        f <- lm.wfit(x,y,d)
    }
    list(coef = f$coef, steps = its)
}
```

With this algorithm, it is straightforward to compute the MR filter at the median:

```
mrfilt <- function (y, theta = 40)
{
    n <- length(y)
    V <- rep(0, n)
    V[1] <- 1
    V[2] <- -2
    V[3] <- 1
    T <- toeplitz(V)
    T[lower.tri(T)] <- 0
    m <- n-2
    D <- T[1:m,]
    X <- rbind(diag(n), theta * D)
    y_tilde <- c(y, rep(0,m))
    out <- meketon(y_tilde,X)
    return(out$coef)
}
```

```
mrt4 <- mrfilt(y)

write.csv(mrt4,"mr_out3.csv")

qplot(y = mrt4, x = time, geom = "path") +
labs(x = "Time", y = "100 x log(GDP)" ) + theme_bw()
```



6 Computing the HP filter with OLS

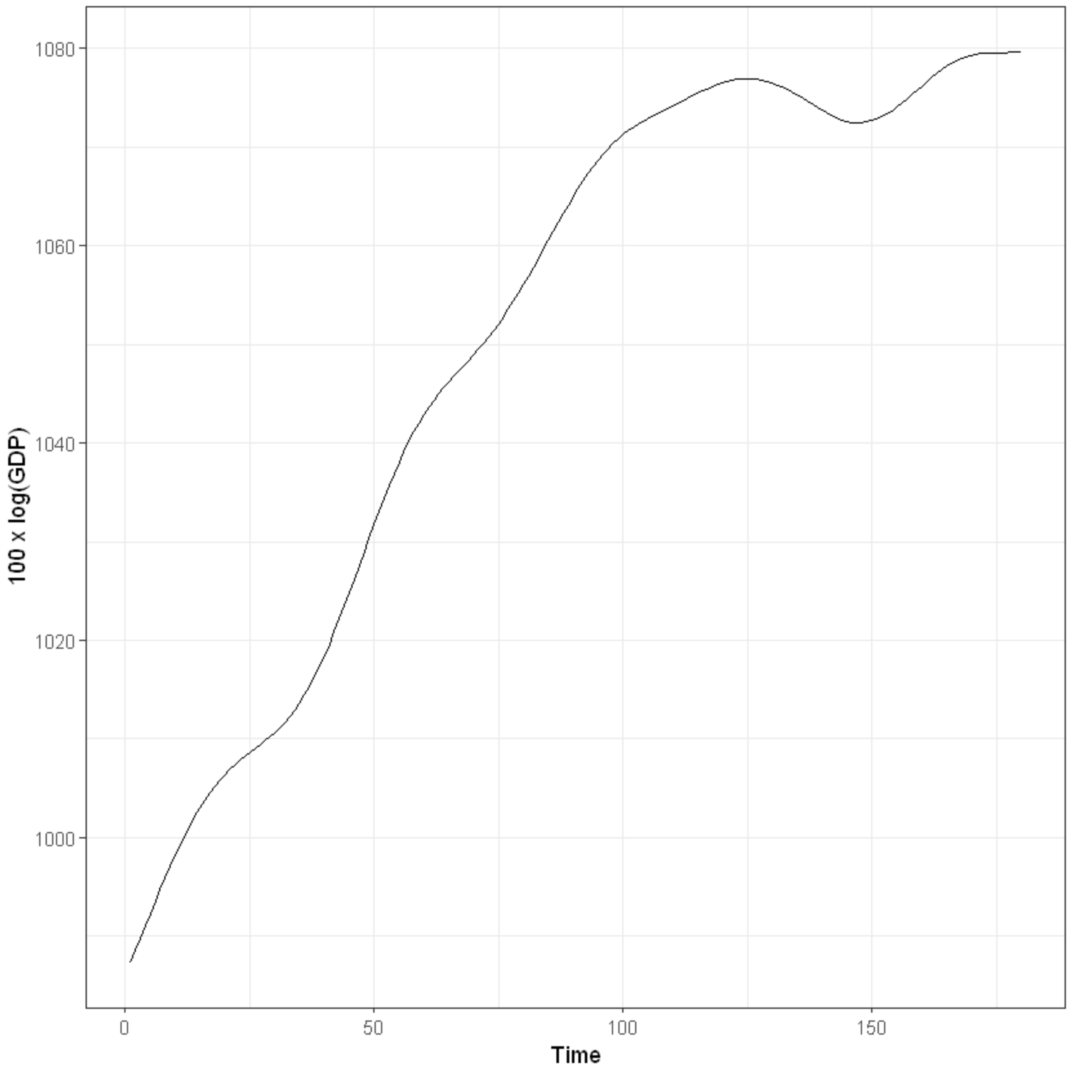
A simple regression (OLS) of $\tilde{y}$ on X, as defined above, produces the Hodrick-Prescott (HP) filter when $\theta = \sqrt{\lambda}$, noting that the HP and MR objetive functions are similar, but the HP uses the ℓ_2-norm:

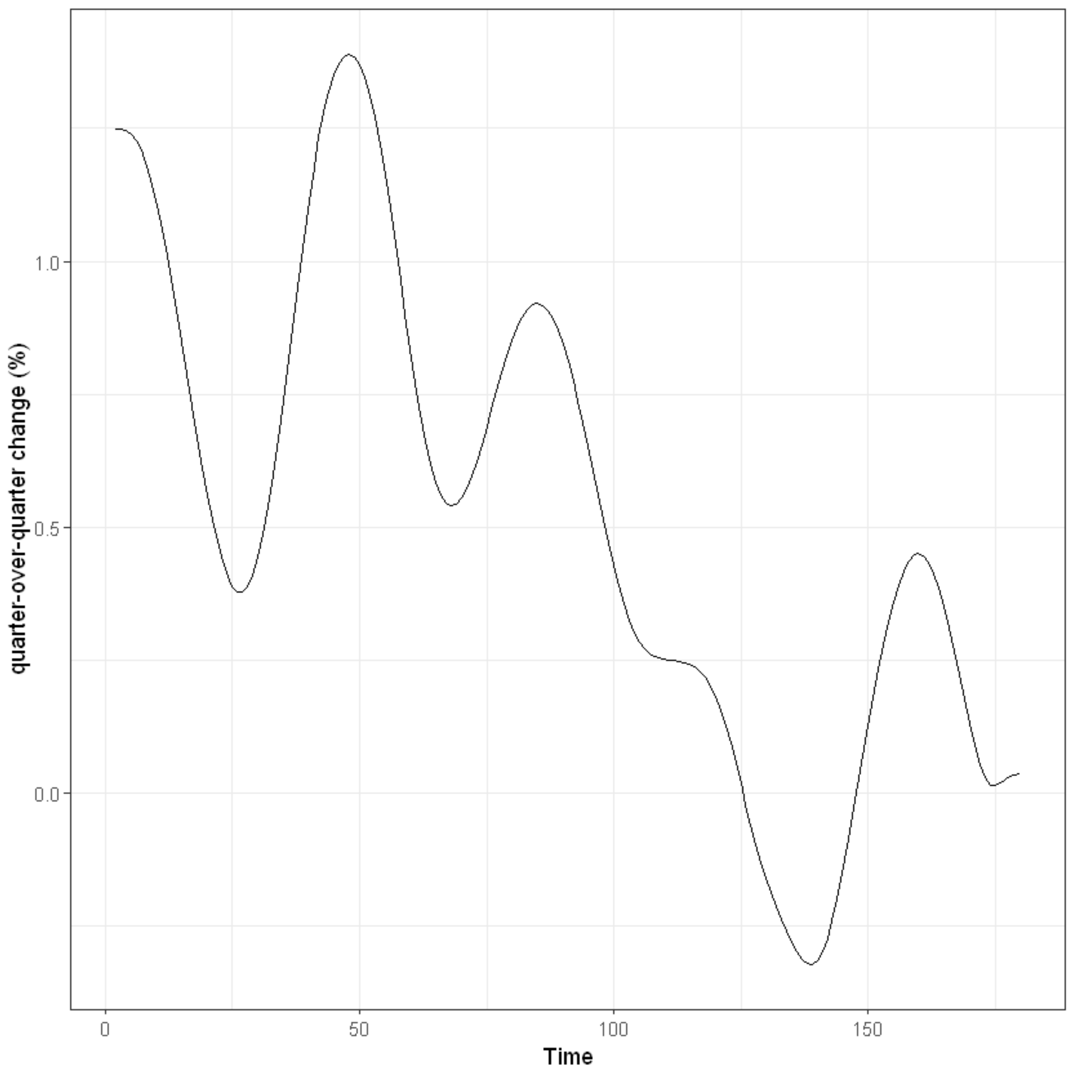
$$\min \left\{ \sum_{t=1}^{N} (y_t - h_t)^2 + \lambda \sum_{t=1}^{N-2} [(h_{t+2} - h_{t+1}) - (h_{t+1} - h_t)]^2 \right\} \tag{12}$$

The code is straightforward in this case, using the usual linear methods lm(y ~ x, data) function from R:

```
mod4 <- lm(model, data = dt)
hpt <- coef(mod4)
dhpt <- hpt[2:n] - hpt[1:(n-1)]

write.csv(hpt,"hp_out.csv")
```

```
qplot(y = hpt, x = time, geom = "path") +
labs(x = "Time", y = "100 x log(GDP)" ) + theme_bw()

qplot(y = dhpt, x = time[2:n], geom = "path") +
labs(x = "Time", y = "quarter-over-quarter change (%)" ) + theme_bw()
```

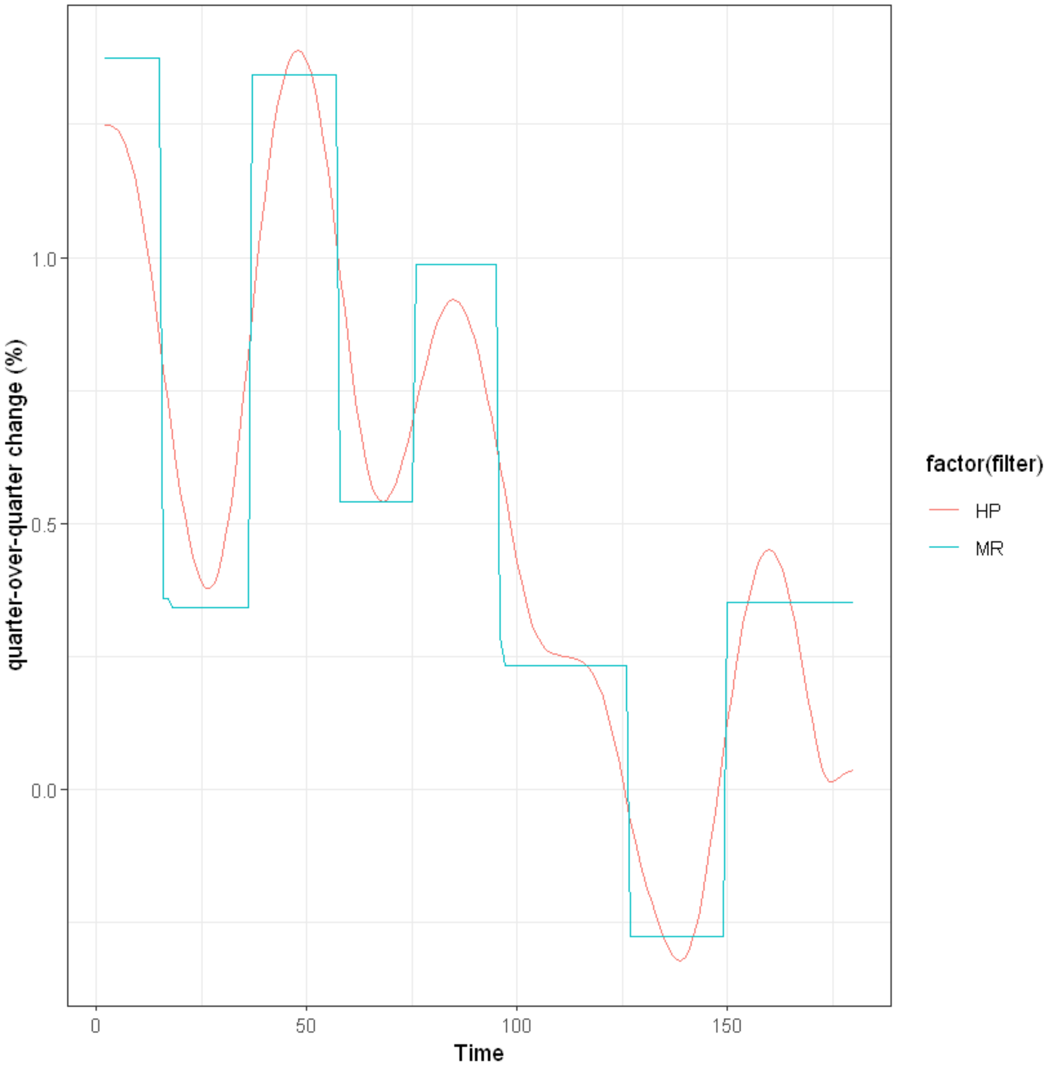




7 Conclusion

The figure below compares the first differences of the HP and MR trends computed with OLS and LAD, respectively.

```
dtrend <- c(dhpt,dmrt1)
filter <- c(rep("HP",(n-1)),rep("MR",(n-1)))
time4 <- c(time[2:n],time[2:n])
qplot(y = dtrend, x = time4, geom = "path", col = factor(filter)) +
labs(x = "Time", y = "quarter-over-quarter change (%)") + theme_bw()
```



It is evident that the MR filter produces a more stable and robust trend than the HP filter with a similar computing effort. In particular, the MR filter deals in a nice way with the structural change concerned with COVID-19 pandemic at the end of the sample, avoiding the oscillation produced by the HP and maintaining a quarter-over-quarter potential GDP growth around 0.35% or 1.4% annualized.

Acknowledgments

This work was supported by Fundação para a Ciência e Tecnologia (FCT), Lisbon, Portugal under post-doctoral grant with the reference CUBE-MACROECO-BPD4.

References

Hodrick, R. J., & Prescott, E. C. (1997). Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking*, 29(1), 1–16.

Kim, S.-J., Koh, K., Boyd, S., & Gorinevsky, D. (2009). ℓ_1 Trend Filtering. *Society for Industrial and Applied Mathematics (SIAM) Review*, 51(2), 339–360.

Koenker, R. (2008). Quantile Regression Computation: From the Inside and the Outside. Retrieved from http://www.econ.uiuc.edu/~roger/research/rq/QRComputation.pdf

Koenker, R. (2019). Quantile Regression in R: a vignette. Retrieved from https://cran.r-project.org/

Meketon, M. S. (1986). Least absolute value regression. Bell Labs Technical Report.

Mosheiov, G., & Raveh, A. (1997). On trend estimation of time-series: a simple linear programming approach. *Journal of the Operational Research Society*, 48, 90–96.

Portnoy, S., & Koenker, R. (1997). The Guassian Hare and the Laplacian Tortoise: Computability of Squared-Error versus Absolute-Error Estimators. *Statistical Science*, 12(4), 279–296.

Ravn, M. O., & Uhlig, H. (2002). Notes on adjusting the Hodrick-Prescott filter for the frequency of observations. *The Review of Economics and Statistics*, (84(2)), 371–380.