

## week12: Die Maximum Likelihood Methode, Teil2

Wir wollen uns jetzt noch das Beispiel 3 vom letzten Mal anschauen, das war die folgende Sache:

- In Beispiel 3 sind die Returns einer Finanzzeitreihe gegeben und wir möchten einen Zeithorizont  $d$  so wählen, dass die mit einer  $d$ -Tages-Volatilität normierten Returns möglichst gut durch eine Standard-Normalverteilung approximiert werden können. Wie können wir das optimale  $d$  bestimmen?

Erinnern wir uns zunächst an die Definition der normierten Returns wie wir sie im week2.pdf im letzten Beispiel 3 angegeben hatten:

Es bezeichne

$$\{ S(t_k) \}_{k=0}^n \quad (1)$$

die Zeitreihe eines liquide handelbaren Assets, etwa die täglichen Schlusskurse des S&P500 Aktienindex, der umfasst die 500 grössten börsennotierten Unternehmen in den USA und ist ein standard Aktienindex für ökonometrische Analysen. Die täglichen Returns am Tag  $t_k$  sind definiert durch 'Preis heute minus Preis gestern durch Preis gestern', also genauer:

$$\text{ret}(t_k) := \frac{S(t_k) - S(t_{k-1})}{S(t_{k-1})} \quad (2)$$

Mit Hilfe dieser Returns kann man für einen gegebenen Tag  $t_k$  mit  $k \geq d$  eine  $d$ -Tages-Volatilität  $\text{vol}_d(t_k)$  berechnen. Diese ist definiert durch die Formel

$$\text{vol}_d(t_k) := \left\{ \frac{1}{d} \sum_{j=0}^{d-1} [\text{ret}(t_{k-j})]^2 \right\}^{1/2} \quad (3)$$

Typische Werte für  $d$  sind etwa  $d = 15$  Tage oder  $d = 20$  Tage, man berücksichtigt also jeweils die letzten 3 oder 4 Wochen, eine Woche hat 5 Handelstage. Definiert man nun normierte oder standardisierte Returns durch

$$\text{normret}_d(t_k) := \frac{\text{ret}(t_k)}{\text{vol}_d(t_{k-1})} \quad (4)$$

dann stellt man fest (das hatten wir konkret im Kapitel 6 gemacht), dass sich diese normierten Returns recht gut durch eine Standard-Normalverteilung approximieren lassen, es gilt die folgende approximative Identität:

$$\text{Prob} \left[ \text{normret}_d(t_k) \in (x, x + dx) \right] \approx e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}} \quad (5)$$

Dabei hatten wir gesehen, dass das für manche  $d$ 's besser hingekommen ist und für manche  $d$ 's nicht so gut, das kann man sich auch nochmal auf dem Bild in der ersten Aufgabe vom Ü-Blatt 6 anschauen. Hier möchten wir jetzt also mit der Maximum-Likelihood-Methode das optimale  $d$  bestimmen. Dazu müssen wir uns zunächst an den Begriff der bedingten Wahrscheinlichkeit erinnern: Es seien  $A$  und  $B$  zwei Ereignisse, etwa

- $A$  : wir würfeln eine 4  
 $B$  : wir würfeln eine gerade Zahl

Dann ist die bedingte Wahrscheinlichkeit  $P(A|B)$ , die W'keit, dass das Ereignis  $A$  eintritt unter der Voraussetzung, dass das Ereignis  $B$  schon eingetreten ist, definiert durch

$$P(A|B) := \frac{P(A \cap B)}{P(B)} \quad (6)$$

In dem konkreten Beispiel mit dem Würfeln wäre das dann also, die W'keit, eine 4 zu würfeln, unter der Voraussetzung, dass wir schon wissen, dass eine gerade Zahl kommt,

$$P(A|B) := \frac{P(\{4\} \cap \{2, 4, 6\})}{P(\{2, 4, 6\})} = \frac{P(\{4\})}{P(\{2, 4, 6\})} = \frac{1/6}{3/6} = \frac{1}{3}$$

Damit können wir jetzt die Likelihood-Funktion aufstellen:

Zunächst mal können wir die Gleichung (5) unter Berücksichtigung von (4) auch folgendermassen schreiben:

$$\text{Prob} \left[ \text{ret}(t_k) \in \text{vol}_d(t_{k-1})(x, x + dx) \right] \approx e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}} \quad (7)$$

Wir setzen

$$\begin{aligned} y &= y_k := \text{vol}_d(t_{k-1})x && \Leftrightarrow && x &= \frac{y_k}{\text{vol}_d(t_{k-1})} \\ dy &= dy_k := \text{vol}_d(t_{k-1})dx && \Leftrightarrow && dx &= \frac{dy_k}{\text{vol}_d(t_{k-1})} \end{aligned}$$

und bekommen dann

$$\text{Prob} \left[ \text{ret}(t_k) \in (y_k, y_k + dy_k) \right] = \exp \left\{ -\frac{y_k^2}{2 \text{vol}_d^2(t_{k-1})} \right\} \frac{dy_k}{\sqrt{2\pi} \text{vol}_d(t_{k-1})} \quad (8)$$

Von jetzt an wollen wir ein exaktes Gleichheitszeichen in (8) schreiben und kein  $\approx$ , weil das jetzt unser statistisches Modell ist: Wenn sich die Returns

$$\text{ret}(t_1), \text{ret}(t_2), \dots, \text{ret}(t_{k-1})$$

bereits alle realisiert haben, so dass die Grösse

$$\text{vol}_d^2(t_{k-1}) = \frac{1}{d} \left[ \text{ret}^2(t_{k-1}) + \text{ret}^2(t_{k-2}) + \dots + \text{ret}^2(t_{k-d}) \right]$$

also eine gegebene, konkrete Zahl ist, dann ist der  $k$ -te Return  $\text{ret}(t_k)$  normalverteilt mit Mittelwert 0 und Standardabweichung  $\sigma = \text{vol}_d(t_{k-1})$ , genau das soll durch die Gleichung (8) ausgedrückt werden. Da wir da also annehmen, dass sich die vergangenen Returns bereits

realisiert haben, ist das genau genommen eine bedingte Wahrscheinlichkeit und wir müssten genauer folgendes schreiben:

$$\begin{aligned} & \text{Prob} \left[ \text{ret}(t_k) \in (y_k, y_k + dy_k) \mid \text{ret}(t_{k-1}), \dots, \text{ret}(t_1) \right] \\ &= \exp \left\{ -\frac{y_k^2}{2 \text{vol}_d^2(t_{k-1})} \right\} \frac{dy_k}{\sqrt{2\pi} \text{vol}_d(t_{k-1})} \end{aligned} \quad (9)$$

Das ist jetzt die W'keit nur für den  $k$ -ten Return, wenn wir alle Returns nehmen, können wir das dann folgendermassen schreiben:

$$\begin{aligned} & \text{Prob} \left[ \text{ret}(t_k) \in (y_k, y_k + dy_k), \dots, \text{ret}(t_1) \in (y_1, y_1 + dy_1) \right] \\ &= \frac{\text{Prob} \left[ \text{ret}(t_k) \in (y_k, y_k + dy_k), \dots, \text{ret}(t_1) \in (y_1, y_1 + dy_1) \right]}{\text{Prob} \left[ \text{ret}(t_{k-1}) \in (y_{k-1}, y_{k-1} + dy_{k-1}), \dots, \text{ret}(t_1) \in (y_1, y_1 + dy_1) \right]} \times \\ & \quad \text{Prob} \left[ \text{ret}(t_{k-1}) \in (y_{k-1}, y_{k-1} + dy_{k-1}), \dots, \text{ret}(t_1) \in (y_1, y_1 + dy_1) \right] \\ &= \text{Prob} \left[ \text{ret}(t_k) \in (y_k, y_k + dy_k) \mid \text{ret}(t_{k-1}), \dots, \text{ret}(t_1) \right] \times \\ & \quad \text{Prob} \left[ \text{ret}(t_{k-1}) \in (y_{k-1}, y_{k-1} + dy_{k-1}), \dots, \text{ret}(t_1) \in (y_1, y_1 + dy_1) \right] \end{aligned} \quad (10)$$

Für den letzten Faktor in (10) können wir das dann nochmal so machen mit  $k \rightarrow k-1$ :

$$\begin{aligned} & \text{Prob} \left[ \text{ret}(t_{k-1}) \in (y_{k-1}, y_{k-1} + dy_{k-1}), \dots, \text{ret}(t_1) \in (y_1, y_1 + dy_1) \right] = \\ &= \text{Prob} \left[ \text{ret}(t_{k-1}) \in (y_{k-1}, y_{k-1} + dy_{k-1}) \mid \text{ret}(t_{k-2}), \dots, \text{ret}(t_1) \right] \times \\ & \quad \text{Prob} \left[ \text{ret}(t_{k-2}) \in (y_{k-2}, y_{k-2} + dy_{k-2}), \dots, \text{ret}(t_1) \in (y_1, y_1 + dy_1) \right] \end{aligned} \quad (11)$$

Wir setzen (11) in (10) ein und bekommen

$$\begin{aligned} & \text{Prob} \left[ \text{ret}(t_k) \in (y_k, y_k + dy_k), \dots, \text{ret}(t_1) \in (y_1, y_1 + dy_1) \right] \\ &= \text{Prob} \left[ \text{ret}(t_k) \in (y_k, y_k + dy_k) \mid \text{ret}(t_{k-1}), \dots, \text{ret}(t_1) \right] \times \\ & \quad \text{Prob} \left[ \text{ret}(t_{k-1}) \in (y_{k-1}, y_{k-1} + dy_{k-1}) \mid \text{ret}(t_{k-2}), \dots, \text{ret}(t_1) \right] \times \\ & \quad \text{Prob} \left[ \text{ret}(t_{k-2}) \in (y_{k-2}, y_{k-2} + dy_{k-2}), \dots, \text{ret}(t_1) \in (y_1, y_1 + dy_1) \right] \\ & \stackrel{(9)}{=} \exp \left\{ -\frac{y_k^2}{2 \text{vol}_d^2(t_{k-1})} \right\} \frac{dy_k}{\sqrt{2\pi} \text{vol}_d(t_{k-1})} \times \exp \left\{ -\frac{y_{k-1}^2}{2 \text{vol}_d^2(t_{k-2})} \right\} \frac{dy_{k-1}}{\sqrt{2\pi} \text{vol}_d(t_{k-2})} \\ & \quad \times \text{Prob} \left[ \text{ret}(t_{k-2}) \in (y_{k-2}, y_{k-2} + dy_{k-2}), \dots, \text{ret}(t_1) \in (y_1, y_1 + dy_1) \right] \end{aligned} \quad (12)$$

Insgesamt tun wir das dann  $k - d_{max}$  oder  $k - d_0$  mal iterieren, so dass am Ende nur noch die Returns  $ret(t_1)$  bis  $ret(t_{d_0})$  übrig bleiben und erhalten dann:

$$\begin{aligned} & \text{Prob} \left[ ret(t_k) \in (y_k, y_k + dy_k), \dots, ret(t_1) \in (y_1, y_1 + dy_1) \right] \\ &= \prod_{j=1}^{k-d_0} \exp \left\{ -\frac{y_{k-j+1}^2}{2 \text{vol}_d^2(t_{k-j})} \right\} \frac{dy_{k-j+1}}{\sqrt{2\pi} \text{vol}_d(t_{k-j})} \times \\ & \quad \text{Prob} \left[ ret(t_{d_0}) \in (y_{d_0}, y_{d_0} + dy_{d_0}), \dots, ret(t_1) \in (y_1, y_1 + dy_1) \right] \end{aligned} \quad (13)$$

Dabei ist das  $d_{max}$  oder das  $d_0$  das grösste mögliche  $d$  was wir überhaupt zulassen wollen, etwa  $d_{max} = 100$ . Die ersten  $d$  Returns brauchen wir ja jeweils, um überhaupt eine  $d$ -Tages-Volatilität  $\text{vol}_d$  berechnen zu können, wir betrachten sie also als gegeben und die entsprechende W'keit, der letzte Faktor in (13), betrachten wir als eine gegebene, fest Zahl (mit  $1 \leq d \leq d_0 = d_{max}$ )

$$C := \text{Prob} \left[ ret(t_{d_0}) \in (y_{d_0}, y_{d_0} + dy_{d_0}), \dots, ret(t_1) \in (y_1, y_1 + dy_1) \right] \quad (14)$$

Die Likelihood-Funktion lautet dann:

$$L = L(d) = \prod_{j=1}^{k-d_0} \exp \left\{ -\frac{y_{k-j+1}^2}{2 \text{vol}_d^2(t_{k-j})} \right\} \frac{dy_{k-j+1}}{\sqrt{2\pi} \text{vol}_d(t_{k-j})} \times C$$

oder, wenn wir die gesamte Länge unserer Finanzzeitreihe vielleicht, wie in (1), mit  $n$  bezeichnen (etwa  $n \approx 16500$  für die SPX-Zeitreihe), dann

$$\begin{aligned} L(d) &= \prod_{j=1}^{n-d_0} \exp \left\{ -\frac{y_{n-j+1}^2}{2 \text{vol}_d^2(t_{n-j})} \right\} \frac{dy_{n-j+1}}{\sqrt{2\pi} \text{vol}_d(t_{n-j})} \times C \\ &\stackrel{k=n-j+1}{=} \prod_{k=d_0+1}^n \exp \left\{ -\frac{y_k^2}{2 \text{vol}_d^2(t_{k-1})} \right\} \frac{dy_k}{\sqrt{2\pi} \text{vol}_d(t_{k-1})} \times C \end{aligned} \quad (15)$$

Dabei sind auf der rechten Seite von (15) für die  $y_k$  die tatsächlich realisierten Returns  $ret(t_k)$  einzusetzen und für das  $\text{vol}_d$  jeweils der Ausdruck (3). Um das Maximum zu finden, nehmen wir wieder den Logarithmus:

$$\begin{aligned} \log L(d) &= - \sum_{k=d_0+1}^n \frac{y_k^2}{2 \text{vol}_d^2(t_{k-1})} + \sum_{k=d_0+1}^n \log \left[ \frac{dy_k}{\sqrt{2\pi} \text{vol}_d(t_{k-1})} \right] + \log C \\ &= - \sum_{k=d_0+1}^n \frac{y_k^2}{2 \text{vol}_d^2(t_{k-1})} - \sum_{k=d_0+1}^n \log [\text{vol}_d(t_{k-1})] + \text{const} \end{aligned} \quad (16)$$

mit einer von  $d$  unabhängigen Konstanten ( $d_0 = 100$  ist fest und das  $d$  variiert zwischen 1 und  $d_0 = 100$ )

$$\text{const} := \sum_{k=d_0+1}^n \log \left[ \frac{dy_k}{\sqrt{2\pi}} \right] + \log C$$

Wir müssen also das Maximum von folgender Funktion bestimmen (mit  $d_0 := 100$ ), indem wir das  $d$  alle ganzen Zahlen zwischen 1 und 100 durchlaufen lassen:

$$F(d) = F(d \mid \{y_k\}) := -\frac{1}{2} \sum_{k=101}^n \left\{ \frac{y_k^2}{\text{vol}_d^2(t_{k-1})} + \log[\text{vol}_d^2(t_{k-1})] \right\} \quad (17)$$

mit

$$\begin{aligned} \text{vol}_d^2(t_{k-1}) &= \frac{1}{d} \left[ \text{ret}^2(t_{k-1}) + \text{ret}^2(t_{k-2}) + \dots + \text{ret}^2(t_{k-d}) \right] \\ &= \frac{1}{d} \left[ y_{k-1}^2 + y_{k-2}^2 + \dots + y_{k-d}^2 \right] \end{aligned} \quad (18)$$

Das können wir in diesem Fall nur noch numerisch machen.

→ Start R-Session

**Bemerkung:** Wir haben ein bisschen was verschwiegen, das wäre die folgende Sache: Das statistische Modell

$$\text{Prob} \left[ \text{normret}_d(t_k) \in (x, x + dx) \right] = e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}}$$

ist äquivalent zu

$$\text{normret}_d(t_k) = \frac{\text{ret}(t_k)}{\text{vol}_d(t_{k-1})} = \phi_k$$

oder

$$\text{ret}(t_k) = \frac{S(t_k) - S(t_{k-1})}{S(t_{k-1})} = \text{vol}_d(t_{k-1}) \phi_k \quad (19)$$

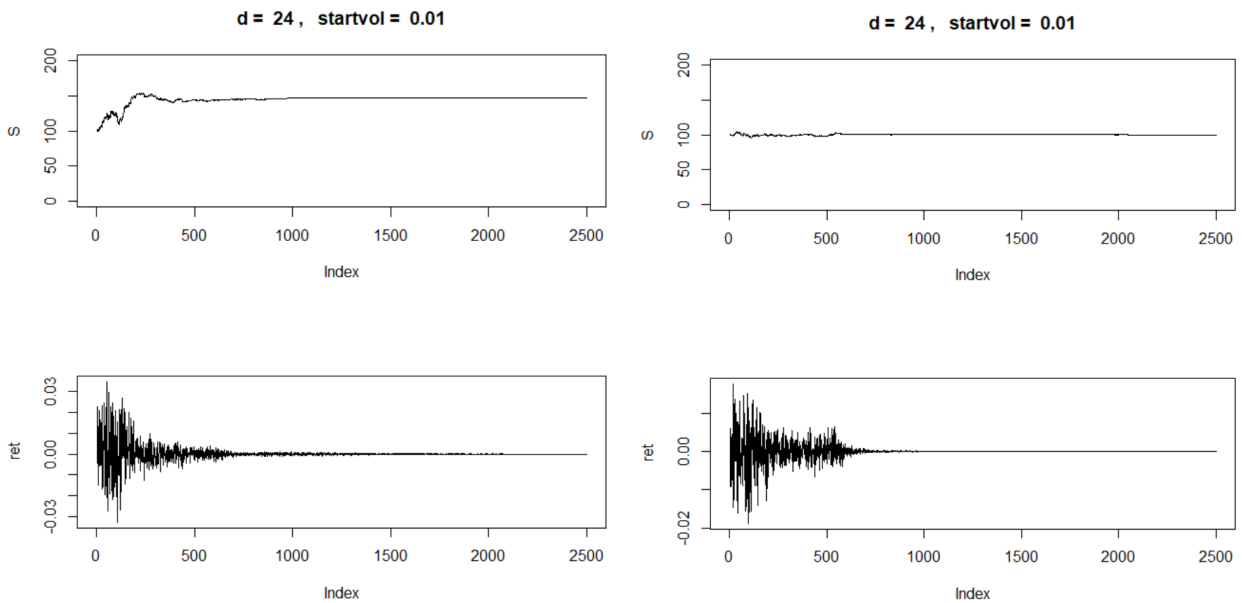
wobei  $\phi_k$  eine standard-normalverteilte Zufallszahl ist. Die Gleichung (19) können wir nach  $S(t_k)$  auflösen, wir bekommen

$$S(t_k) = S(t_{k-1}) \left[ 1 + \text{vol}_d(t_{k-1}) \phi_k \right] \quad (20)$$

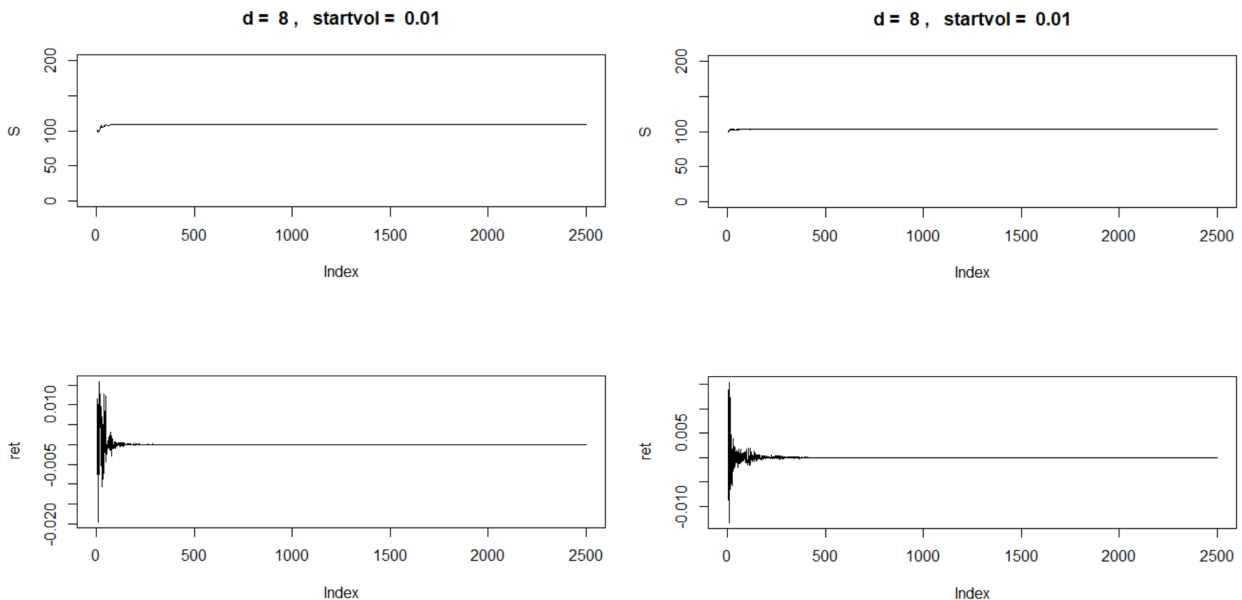
mit der d-Tages-Volatilität

$$\text{vol}_d^2(t_{k-1}) = \frac{1}{d} \left[ \text{ret}^2(t_{k-1}) + \text{ret}^2(t_{k-2}) + \dots + \text{ret}^2(t_{k-d}) \right] \quad (21)$$

Die Gleichung (20) ist eine stochastische Rekursion für die  $S(t_k)$ , aus den gegebenen Daten bis Zeit  $t_{k-1}$  und dem Ziehen einer Zufallszahl  $\phi_k$  können wir einen neuen Wert  $S(t_k)$  berechnen. Wenn man das macht, mit der vol-Spezifikation (21) und einem  $d$  von, nehmen wir mal das optimale  $d$  für die SPX-Zeitreihe, wir hatten den Wert von  $d = 24$  gefunden, bekommt man etwa die folgenden Bilder:



Wenn man das  $d$  kleiner macht, wird das noch extremer, etwa für  $d = 8$ :



Offensichtlich liefert die Modell-Spezifikation (20,21) kein vernünftiges stochastisches Modell. Das kann man jetzt natürlich noch etwas genauer analysieren, woran das denn liegt, das würde man dann vielleicht in einer Spezialvorlesung über Ökonometrie machen, Tatsache ist, dass wenn man die Modell-Spezifikation (20,21) abändern tut zu

$$S(t_k) = S(t_{k-1}) \left[ 1 + \text{vol}_d(t_{k-1}) \phi_k \right] \quad (22)$$

das ist exakt dasselbe wie (20), wir müssen nur die vol-Spezifikation ändern, zu

$$\text{vol}_d^2(t_{k-1}) := w_0 \text{bsvol}^2 + (1 - w_0) \frac{1}{d} \left[ \text{ret}^2(t_{k-1}) + \text{ret}^2(t_{k-2}) + \dots + \text{ret}^2(t_{k-d}) \right] \quad (23)$$

dann bekommt man ‘vernünftige’ Preispfade heraus. Dabei ist das  $w_0$  ein Gewichtsparameter,  $w_0 \in (0, 1)$ , und  $\text{bsvol}$ , das steht etwa für ‘Black-Scholes Volatilität’, ist eine konstante

Volatilität, eine konstant Zahl. Typischerweise, wenn man sie mit Maximum-Likelihood-Estimation bestimmt, findet man

$$\text{bsvol}^2 \approx \frac{1}{n} \sum_{k=1}^n \text{ret}^2(t_k) \stackrel{\text{SPX-Zeitreihe}}{\approx} 1\% = 0.01 \quad (24)$$

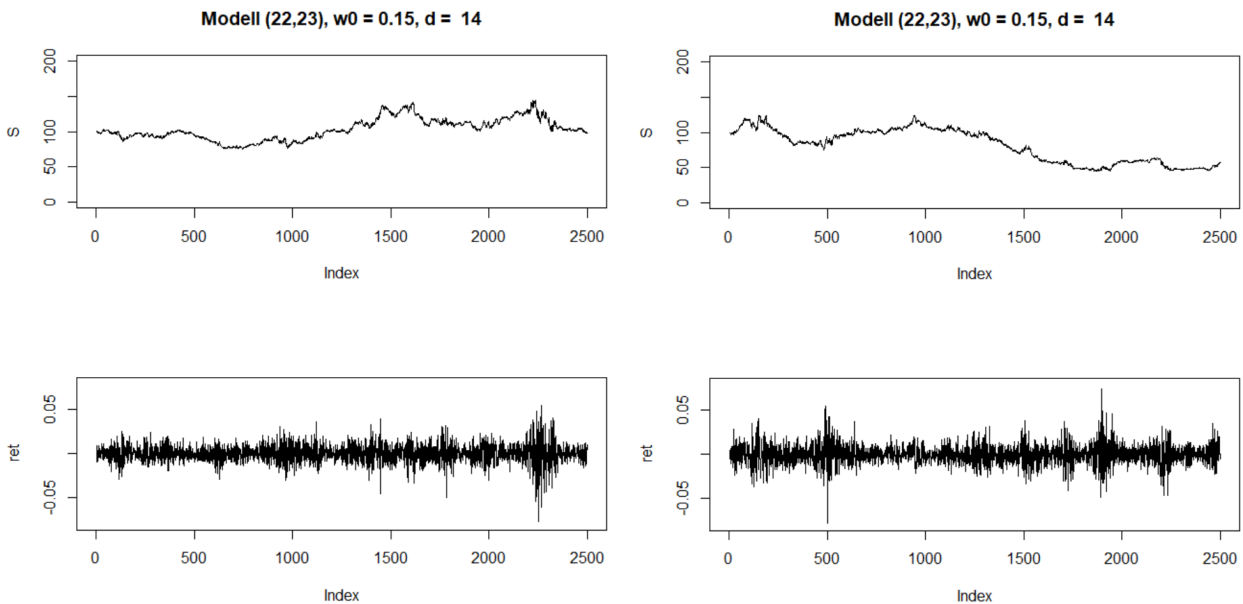
Ein typischer Wert für das  $w_0$  ist

$$w_0 \approx 0.15 \quad (25)$$

und Maximum-Likelihood Schätzung für das  $d$  liefert für den SPX dann den Wert

$$d = 14 \quad (26)$$

Wenn man dann mit diesen Werten (24-26) Pfade für das neue Modell (22,23) simuliert, sehen die typischerweise dann folgendermassen aus:



Dann sieht dann also nicht unplausibel aus.