GARCH Option Pricing[a]

- Options can be priced when the underlying asset's return follows a GARCH process.
- Let S_t denote the asset price at date t.
- Let h_t^2 be the conditional variance of the return over the period $[\,t, t+1\,]$ given the information at date t.
  - "One day" is merely a convenient term for any elapsed time Δt.

[a] A Bloomberg quant said, on Feb 29, 2008, that GARCH option pricing is seldom used in trading.

GARCH Option Pricing (continued)

- Adopt the following risk-neutral process for the price dynamics:[a]

$$\ln \frac{S_{t+1}}{S_t} = r - \frac{h_t^2}{2} + h_t \epsilon_{t+1}, \tag{66}$$

where

$$\begin{aligned} h_{t+1}^2 &= \beta_0 + \beta_1 h_t^2 + \beta_2 h_t^2 (\epsilon_{t+1} - c)^2, \qquad (67) \\ \epsilon_{t+1} &\sim N(0,1) \text{ given information at date } t, \\ r &= \text{daily riskless return}, \\ c &\geq 0. \end{aligned}$$

[a] Duan (1995).

GARCH Option Pricing (continued)

- The five unknown parameters of the model are c, h_0, β_0, β_1, and β_2.
- It is postulated that $\beta_0, \beta_1, \beta_2 \geq 0$ to make the conditional variance positive.
- The above process, called the nonlinear asymmetric GARCH model, generalizes the GARCH$(1, 1)$ model (see text).

GARCH Option Pricing (continued)

- It captures the volatility clustering in asset returns first noted by Mandelbrot (1963).[a]
  - When $c = 0$, a large ϵ_{t+1} results in a large h_{t+1}, which in turns tends to yield a large h_{t+2}, and so on.
- It also captures the negative correlation between the asset return and changes in its (conditional) volatility.[b]
  - For $c > 0$, a positive ϵ_{t+1} (good news) tends to decrease h_{t+1}, whereas a negative ϵ_{t+1} (bad news) tends to do the opposite.

---

[a] "... large changes tend to be followed by large changes—of either sign—and small changes tend to be followed by small changes ..."

[b] Noted by Black (1976): Volatility tends to rise in response to "bad news" and fall in response to "good news."

GARCH Option Pricing (concluded)

- With $y_t \equiv \ln S_t$ denoting the logarithmic price, the model becomes

$$y_{t+1} = y_t + r - \frac{h_t^2}{2} + h_t \epsilon_{t+1}. \tag{68}$$

- The pair (y_t, h_t^2) completely describes the current state.
- The conditional mean and variance of y_{t+1} are clearly

$$E[\, y_{t+1} \,|\, y_t, h_t^2 \,] = y_t + r - \frac{h_t^2}{2}, \tag{69}$$

$$\mathrm{Var}[\, y_{t+1} \,|\, y_t, h_t^2 \,] = h_t^2. \tag{70}$$

The Ritchken-Trevor (RT) Algorithm[a]

- The GARCH model is a continuous-state model.
- To approximate it, we turn to trees with *discrete* states.
- Path dependence in GARCH makes the tree for asset prices explode exponentially (why?).
- We need to mitigate this combinatorial explosion.

[a] Ritchken and Trevor (1999).

The Ritchken-Trevor Algorithm (continued)

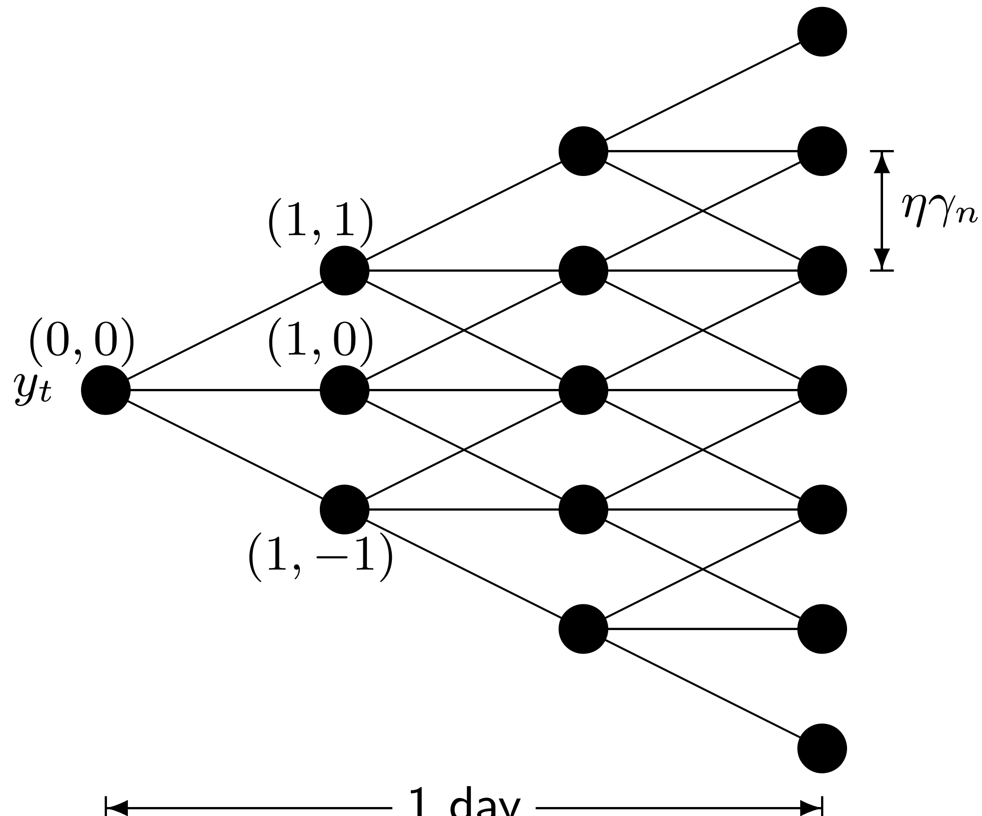
- Partition a day into n periods.
- Three states follow each state (y_t, h_t^2) after a period.
- As the trinomial model combines, $2n+1$ states at date $t+1$ follow each state at date t (recall p. 549).
- These $2n+1$ values must approximate the distribution of (y_{t+1}, h_{t+1}^2).
- So the conditional moments (69)–(70) at date $t+1$ on p. 690 must be matched by the trinomial model to guarantee convergence to the continuous-state model.

The Ritchken-Trevor Algorithm (continued)

- It remains to pick the jump size and the three branching probabilities.
- The role of σ in the Black-Scholes option pricing model is played by h_t in the GARCH model.
- As a jump size proportional to $\sigma/\sqrt{n}$ is picked in the BOPM, a comparable magnitude will be chosen here.
- Define $\gamma \equiv h_0$, though other multiples of h_0 are possible, and

$$\gamma_n \equiv \frac{\gamma}{\sqrt{n}}.$$

- The jump size will be some integer multiple η of γ_n.
- We call η the jump parameter (p. 694).



The seven values on the right approximate the distribution of logarithmic price y_{t+1}.

The Ritchken-Trevor Algorithm (continued)

- The middle branch does not change the underlying asset's price.
- The probabilities for the up, middle, and down branches are

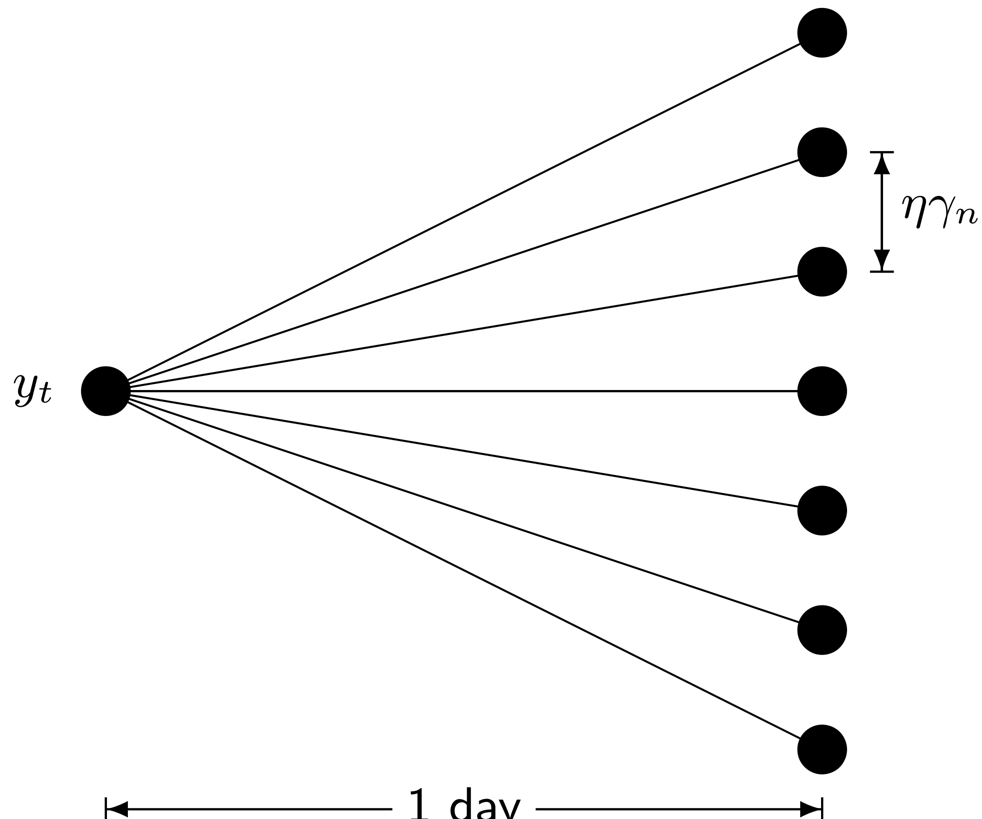
$$
\begin{aligned}
p_u &= \frac{h_t^2}{2\eta^2\gamma^2} + \frac{r - (h_t^2/2)}{2\eta\gamma\sqrt{n}}, && (71)\\
p_m &= 1 - \frac{h_t^2}{\eta^2\gamma^2}, && (72)\\
p_d &= \frac{h_t^2}{2\eta^2\gamma^2} - \frac{r - (h_t^2/2)}{2\eta\gamma\sqrt{n}}. && (73)
\end{aligned}
$$

The Ritchken-Trevor Algorithm (continued)

- It can be shown that:
  - The trinomial model takes on $2n+1$ values at date $t+1$ for y_{t+1}.
  - These values have a matching mean for y_{t+1}.
  - These values have an asymptotically matching variance for y_{t+1}.
- The central limit theorem thus guarantees the desired convergence as n increases.

The Ritchken-Trevor Algorithm (continued)

- We can dispense with the intermediate nodes *between* dates to create a $(2n+1)$-nomial tree (p. 698).
- The resulting model is multinomial with $2n+1$ branches from any state (y_t, h_t^2).
- There are two reasons behind this manipulation.
  - Interdate nodes are created merely to approximate the continuous-state model after one day.
  - Keeping the interdate nodes results in a tree that can be as much as n times larger.



This heptanomial tree is the outcome of the trinomial tree on p. 694 after its intermediate nodes are removed.

The Ritchken-Trevor Algorithm (continued)

- A node with logarithmic price $y_t + \ell\eta\gamma_n$ at date $t+1$ follows the current node at date t with price y_t for some $-n \leq \ell \leq n$.
- To reach that price in n periods, the number of up moves must exceed that of down moves by exactly ℓ.
- The probability that this happens is

$$P(\ell) \equiv \sum_{j_u, j_m, j_d} \frac{n!}{j_u!\, j_m!\, j_d!}\, p_u^{j_u} p_m^{j_m} p_d^{j_d},$$

with $j_u, j_m, j_d \geq 0$, $n = j_u + j_m + j_d$, and $\ell = j_u - j_d$.

The Ritchken-Trevor Algorithm (continued)

- A particularly simple way to calculate the $P(\ell)$s starts by noting that

$$(p_u x + p_m + p_d x^{-1})^n = \sum_{\ell=-n}^{n} P(\ell)\, x^{\ell}. \tag{74}$$

- So we expand $(p_u x + p_m + p_d x^{-1})^n$ and retrieve the probabilities by reading off the coefficients.
- It can be computed in $O(n^2)$ time.

The Ritchken-Trevor Algorithm (continued)

- The updating rule (67) on p. 687 must be modified to account for the adoption of the discrete-state model.
- The logarithmic price $y_t + \ell\eta\gamma_n$ at date $t+1$ following state (y_t, h_t^2) at date t has a variance equal to

$$h_{t+1}^2 = \beta_0 + \beta_1 h_t^2 + \beta_2 h_t^2 (\epsilon'_{t+1} - c)^2, \tag{75}$$

  – Above,

$$\epsilon'_{t+1} = \frac{\ell\eta\gamma_n - (r - h_t^2/2)}{h_t}, \quad \ell = 0, \pm 1, \pm 2, \ldots, \pm n,$$

  is a discrete random variable with $2n+1$ values.

The Ritchken-Trevor Algorithm (continued)

- Different conditional variances h_t^2 may require different η so that the probabilities calculated by Eqs. (71)–(73) on p. 695 lie between 0 and 1.
- This implies varying jump sizes.
- The necessary requirement $p_m \geq 0$ implies $\eta \geq h_t/\gamma$.
- Hence we try

$$\eta = \lceil h_t/\gamma \rceil, \lceil h_t/\gamma \rceil + 1, \lceil h_t/\gamma \rceil + 2, \ldots$$

until valid probabilities are obtained or until their nonexistence is confirmed.

The Ritchken-Trevor Algorithm (continued)

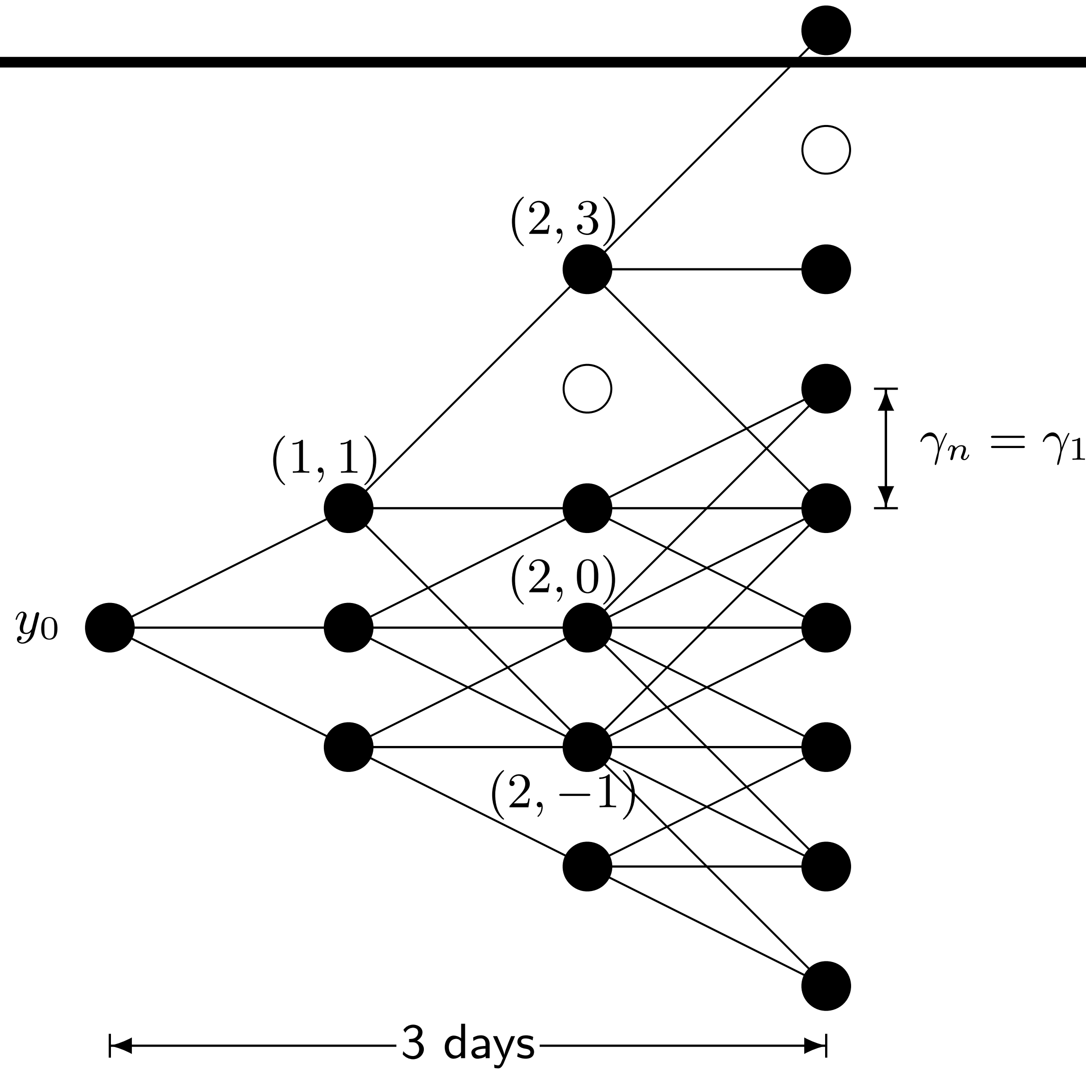
- The sufficient and necessary condition for valid probabilities to exist is[a]

$$\frac{|\, r-(h_t^2/2)\,|}{2\eta\gamma\sqrt{n}} \leq \frac{h_t^2}{2\eta^2\gamma^2} \leq \min\left(1-\frac{|\, r-(h_t^2/2)\,|}{2\eta\gamma\sqrt{n}}, \frac{1}{2}\right).$$

- Obviously, the magnitude of η tends to grow with h_t.
- The plot on p. 704 uses $n=1$ to illustrate our points for a 3-day model.
- For example, node $(1,1)$ of date 1 and node $(2,3)$ of date 2 pick $\eta=2$.

---

[a]Lyuu and Wu (2003).



$(2,3)$
$(1,1)$
$\gamma_n = \gamma_1$
$(2,0)$
y_0
$(2,-1)$
3 days

The Ritchken-Trevor Algorithm (continued)

- The topology of the tree is not a standard combining multinomial tree.
- For example, a few nodes on p. 704 such as nodes $(2, 0)$ and $(2, -1)$ have multiple jump sizes.
- The reason is the path dependence of the model.
  - Two paths can reach node $(2, 0)$ from the root node, each with a different variance for the node.
  - One of the variances results in $\eta = 1$, whereas the other results in $\eta = 2$.

The Ritchken-Trevor Algorithm (concluded)

- The number of possible values of h_t^2 at a node can be exponential.
  - Each path brings with it a different variance h_t^2.
- To address this problem, we record only the maximum and minimum h_t^2 at each node.[a]
- Therefore, each node on the tree contains only two states $(y_t, h_{\max}^2)$ and $(y_t, h_{\min}^2)$.
- Each of $(y_t, h_{\max}^2)$ and $(y_t, h_{\min}^2)$ carries its own η and set of $2n+1$ branching probabilities.

[a]Cakici and Topyan (2000). But see p. 738 for a potential problem.

Negative Aspects of the Ritchken-Trevor Algorithm[a]

- A small n may yield inaccurate option prices.
- But the tree will grow exponentially if n is large enough.
  - Specifically, $n > (1 - \beta_1)/\beta_2$ when $r = c = 0$.
- A large n has another serious problem: The tree cannot grow beyond a certain date.
- Thus the choice of n may be limited in practice.
- The RT algorithm can be modified to be free of shortened maturity and exponential complexity.[b]

---

[a] Lyuu and Wu (2003, 2005).

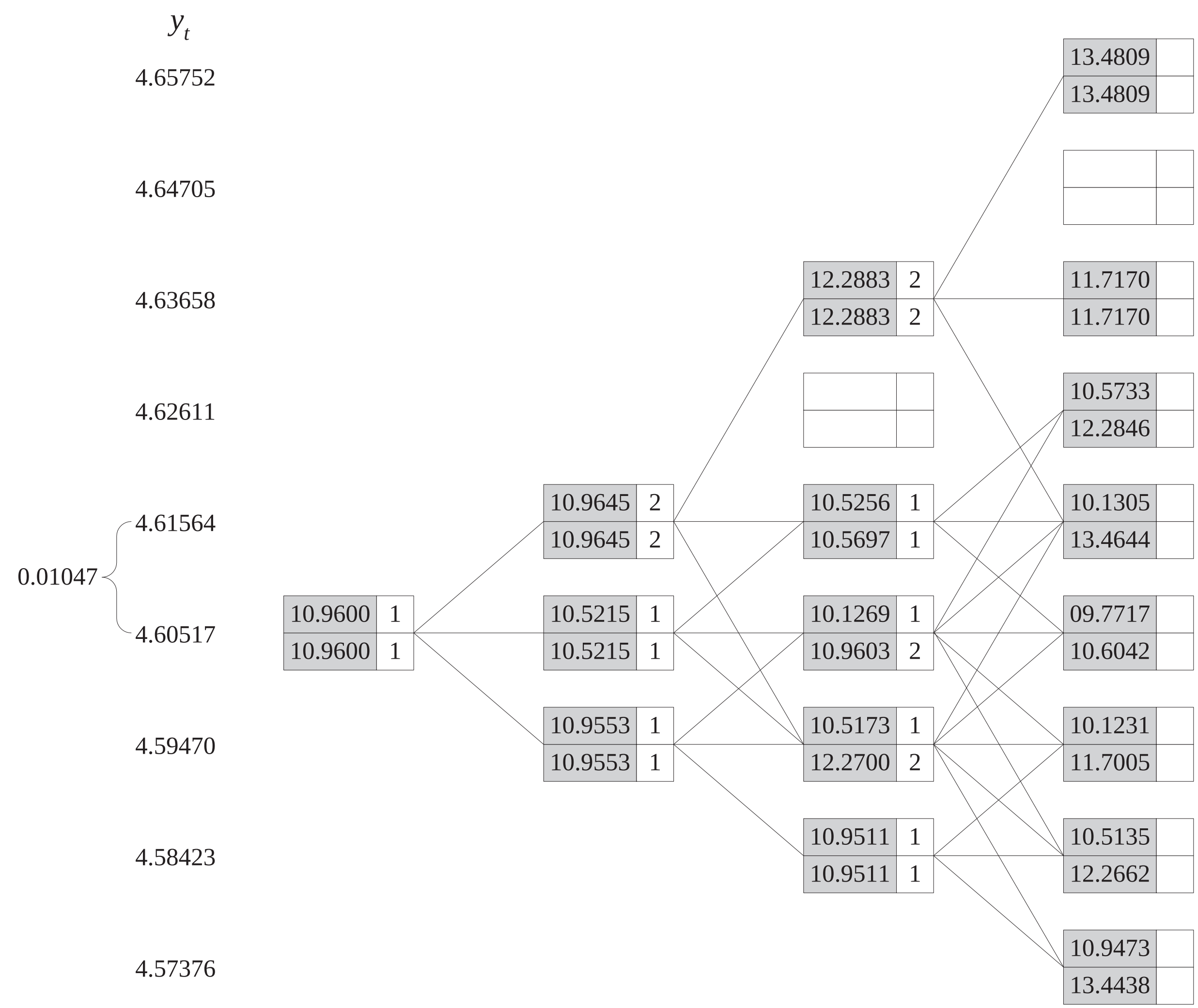
[b] It is only $O(n^2)$ if $n \leq (\sqrt{(1 - \beta_1)/\beta_2} - c)^2$!

Numerical Examples

- Assume $S_0 = 100$, $y_0 = \ln S_0 = 4.60517$, $r = 0$, $h_0^2 = 0.0001096$, $\gamma = h_0 = 0.010469$, $n = 1$, $\gamma_n = \gamma/\sqrt{n} = 0.010469$, $\beta_0 = 0.000006575$, $\beta_1 = 0.9$, $\beta_2 = 0.04$, and $c = 0$.
- A daily variance of 0.0001096 corresponds to an annual volatility of $\sqrt{365 \times 0.0001096} \approx 20\%$.
- Let $h^2(i, j)$ denote the variance at node (i, j).
- Initially, $h^2(0, 0) = h_0^2 = 0.0001096$.

Numerical Examples (continued)

- Let $h^2_{\max}(i, j)$ denote the maximum variance at node (i, j).
- Let $h^2_{\min}(i, j)$ denote the minimum variance at node (i, j).
- Initially, $h^2_{\max}(0, 0) = h^2_{\min}(0, 0) = h^2_0$.
- The resulting three-day tree is depicted on p. 710.


y_t
4.65752
4.64705
4.63658
4.62611
4.61564
0.01047
4.60517
4.59470
4.58423
4.57376
10.9600 1
10.9600 1
10.9645 2
10.9645 2
10.5215 1
10.5215 1
10.9553 1
10.9553 1
12.2883 2
12.2883 2
10.5256 1
10.5697 1
10.1269 1
10.9603 2
10.5173 1
12.2700 2
10.9511 1
10.9511 1
13.4809
13.4809
11.7170
11.7170
10.5733
12.2846
10.1305
13.4644
09.7717
10.6042
10.1231
11.7005
10.5135
12.2662
10.9473
13.4438

A top (bottom) number inside a gray box refers to the minimum (maximum, respectively) variance $h^2_{\min}$ ($h^2_{\max}$, respectively) for the node. Variances are multiplied by 100,000 for readability. A top (bottom) number inside a white box refers to η corresponding to $h^2_{\min}$ ($h^2_{\max}$, respectively).

Numerical Examples (continued)

- Let us see how the numbers are calculated.
- Start with the root node, node $(0, 0)$.
- Try $\eta = 1$ in Eqs. (71)–(73) on p. 695 first to obtain

$$\begin{aligned} p_u &= 0.4974, \\ p_m &= 0, \\ p_d &= 0.5026. \end{aligned}$$

- As they are valid probabilities, the three branches from the root node use single jumps.

Numerical Examples (continued)

- Move on to node $(1, 1)$.
- It has one predecessor node—node $(0, 0)$—and it takes an up move to reach the current node.
- So apply updating rule (75) on p. 701 with $\ell = 1$ and $h_t^2 = h^2(0, 0)$.
- The result is $h^2(1, 1) = 0.000109645$.

Numerical Examples (continued)

- Because $\lceil h(1,1)/\gamma \rceil = 2$, we try $\eta = 2$ in Eqs. (71)–(73) on p. 695 first to obtain

$$\begin{aligned} p_u &= 0.1237, \\ p_m &= 0.7499, \\ p_d &= 0.1264. \end{aligned}$$

- As they are valid probabilities, the three branches from node $(1,1)$ use double jumps.

Numerical Examples (continued)

- Carry out similar calculations for node $(1,0)$ with $\ell = 0$ in updating rule (75) on p. 701.
- Carry out similar calculations for node $(1,-1)$ with $\ell = -1$ in updating rule (75).
- Single jump $\eta = 1$ works in both nodes.
- The resulting variances are

$$\begin{aligned} h^2(1,0) &= 0.000105215, \\ h^2(1,-1) &= 0.000109553. \end{aligned}$$

Numerical Examples (continued)

- Node $(2, 0)$ has 2 predecessor nodes, $(1, 0)$ and $(1, -1)$.
- Both have to be considered in deriving the variances.
- Let us start with node $(1, 0)$.
- Because it takes a middle move to reach the current node, we apply updating rule (75) on p. 701 with $\ell = 0$ and $h_t^2 = h^2(1, 0)$.
- The result is $h_{t+1}^2 = 0.000101269$.

Numerical Examples (continued)

- Now move on to the other predecessor node $(1, -1)$.
- Because it takes an up move to reach the current node, apply updating rule (75) on p. 701 with $\ell = 1$ and $h_t^2 = h^2(1, -1)$.
- The result is $h_{t+1}^2 = 0.000109603$.
- We hence record

$$\begin{aligned} h_{\min}^2(2,0) &= 0.000101269, \\ h_{\max}^2(2,0) &= 0.000109603. \end{aligned}$$

Numerical Examples (continued)

- Consider state $h^2_{\max}(2,0)$ first.
- Because $\lceil h_{\max}(2,0)/\gamma \rceil = 2$, we first try $\eta = 2$ in Eqs. (71)–(73) on p. 695 to obtain

$$\begin{aligned} p_u &= 0.1237, \\ p_m &= 0.7500, \\ p_d &= 0.1263. \end{aligned}$$

- As they are valid probabilities, the three branches from node $(2,0)$ with the maximum variance use double jumps.

Numerical Examples (continued)

- Now consider state $h^2_{\min}(2,0)$.
- Because $\lceil h_{\min}(2,0)/\gamma \rceil = 1$, we first try $\eta = 1$ in Eqs. (71)–(73) on p. 695 to obtain

$$
\begin{aligned}
p_u &= 0.4596,\\
p_m &= 0.0760,\\
p_d &= 0.4644.
\end{aligned}
$$

- As they are valid probabilities, the three branches from node $(2,0)$ with the minimum variance use single jumps.

Numerical Examples (continued)

- Node $(2, -1)$ has 3 predecessor nodes.
- Start with node $(1, 1)$.
- Because it takes a down move to reach the current node, we apply updating rule (75) on p. 701 with $\ell = -1$ and $h_t^2 = h^2(1, 1)$.
- The result is $h_{t+1}^2 = 0.0001227$.

Numerical Examples (continued)

- Now move on to predecessor node $(1, 0)$.
- Because it also takes a down move to reach the current node, we apply updating rule (75) on p. 701 with $\ell = -1$ and $h_t^2 = h^2(1, 0)$.
- The result is $h_{t+1}^2 = 0.000105609$.

Numerical Examples (continued)

- Finally, consider predecessor node $(1, -1)$.
- Because it takes a middle move to reach the current node, we apply updating rule (75) on p. 701 with $\ell = 0$ and $h_t^2 = h^2(1, -1)$.
- The result is $h_{t+1}^2 = 0.000105173$.
- We hence record

$$\begin{aligned} h^2_{\min}(2, -1) &= 0.000105173, \\ h^2_{\max}(2, -1) &= 0.0001227. \end{aligned}$$

Numerical Examples (continued)

- Consider state $h^2_{\max}(2,-1)$.
- Because $\lceil h_{\max}(2,-1)/\gamma \rceil = 2$, we first try $\eta = 2$ in Eqs. (71)–(73) on p. 695 to obtain

$$\begin{aligned} p_u &= 0.1385, \\ p_m &= 0.7201, \\ p_d &= 0.1414. \end{aligned}$$

- As they are valid probabilities, the three branches from node $(2,-1)$ with the maximum variance use double jumps.

Numerical Examples (continued)

- Next, consider state $h^2_{\min}(2, -1)$.
- Because $\lceil h_{\min}(2, -1)/\gamma \rceil = 1$, we first try $\eta = 1$ in Eqs. (71)–(73) on p. 695 to obtain

$$
\begin{aligned}
p_u &= 0.4773, \\
p_m &= 0.0404, \\
p_d &= 0.4823.
\end{aligned}
$$

- As they are valid probabilities, the three branches from node $(2, -1)$ with the minimum variance use single jumps.

Numerical Examples (concluded)

- Other nodes at dates 2 and 3 can be handled similarly.
- In general, if a node has k predecessor nodes, then $2k$ variances will be calculated using the updating rule.
  - This is because each predecessor node keeps two variance numbers.
- But only the maximum and minimum variances will be kept.

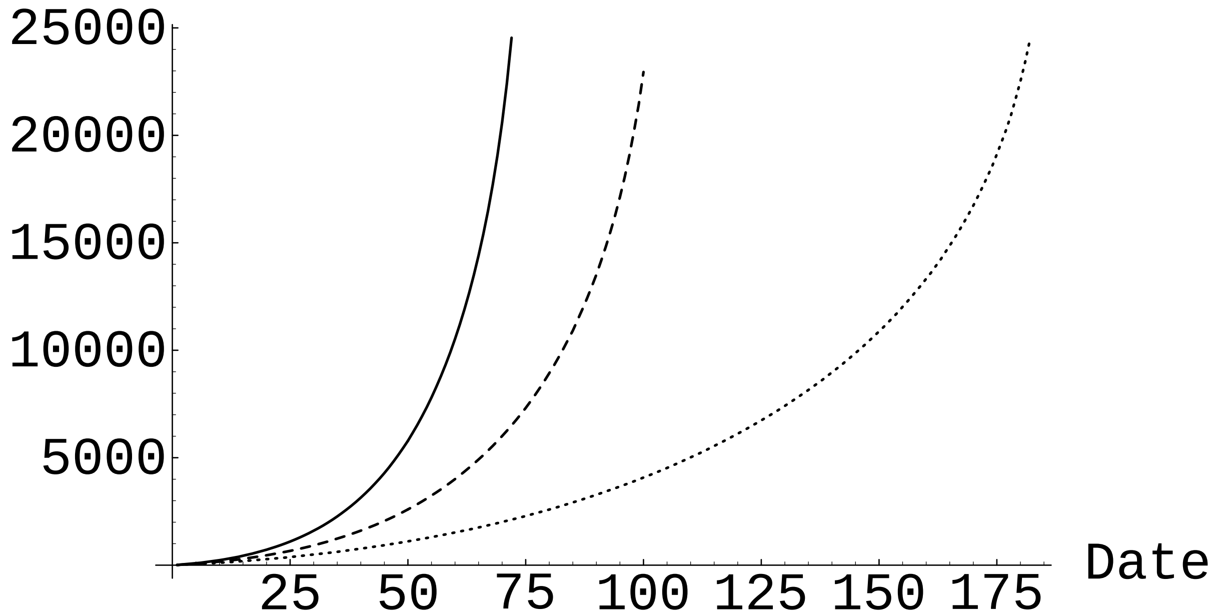
Negative Aspects of the RT Algorithm Revisited[a]

- Recall the problems mentioned on p. 707.
- In our case, combinatorial explosion occurs when

$$n > \frac{1-\beta_1}{\beta_2} = \frac{1-0.9}{0.04} = 2.5.$$

- Suppose we are willing to accept the exponential running time and pick $n = 100$ to seek accuracy.
- But the problem of shortened maturity forces the tree to stop at date 9!

[a]Lyuu and Wu (2003).



Dotted line: $n = 3$; dashed line: $n = 4$; solid line: $n = 5$.