n° 9906

NON–LINEAR INNOVATIONS AND IMPULSE RESPONSES

Christian GOURIEROUX[1]
Joanna JASIAK[2]

This version: July 3, 1999
(First draft: October 26, 1998)

(1) CREST and CEPREMAP, Email: gouriero@ensae.fr
(2) York University, Email: jasiak@yorku.ca

The second author gratefully acknowledges the financial support of the National Sciences and Engineering Research Council of Canada.

Nonlinear Innovations and Impulse Responses

Christian Gourieroux, Joanna Jasiak

Abstract

This paper introduces a concept of innovation for the analysis of nonlinear dynamics. We show that nonlinear processes can be represented as functions of current and lagged values of such innovations. The residuals from nonlinear dynamic models axe used to construct various specification tests. We define and study nonlinear impulse response functions to transitory and permanent shocks.

Innovations non-linéaires et fonctions réponses

Christian Gourieroux, Joanna Jasiak

Résumé

Nous introduisons un concept d'innovation adapté à l'analyse des dynamiques non linéaires. Nous expliquons comment le processus initial peut être exprimé en fonction des valeurs présentes et passées de l'innovation, utilisons les résidus associés pour construire des tests de spécification d'une dynamique non linéaire et pour définir des fonctions réponses à des chocs transitoires ou permanents.

Mots clés : Dynamiques non linéaires, innovation gaussienne, développement de Volterra, fonction réponse, modèle ACD, valeur à risque.

Keywords: Nonlinear Dynamics, Gaussian Innovations, Volterra Expansion, Impulse Response, ACD Model, Value at Risk.

JEL : C5

1 Introduction

The innovations of a stochastic process (Y_t) are usually defined either as a) errors representing differences between the expected and realized values of Y_t, i.e. $\epsilon_t^1 = Y_t - E_{t-1}Y_t$, or b) conditionally standardized expectation errors, i.e. $\epsilon_t^2 = (Y_t - E_{t-1}Y_t)/(V_{t-1}Y_t)^{\frac{1}{2}}$, where $E_{t-1}Y_t$ and $V_{t-1}Y_t$ are the conditional mean and variance of Y_t given the information available at time $t-1$. These definitions cause serious difficulties in the analysis of nonlinear dynamics. For example, the standardized innovations (ϵ_t^2) may not necessarily be independent, due to unobserved cross effects of their conditional moments of order strictly larger than two. Moreover, the definitions of innovations of a series (Y_t) and of its nonlinear transform, such as $(\exp Y_t)$ are not identical. For example, $\epsilon_t^3 = (\exp Y_t - E_{t-1}\exp Y_t)/V_{t-1}(\exp Y_t)$ and ϵ_t^2 do not even satisfy a one-to-one relationship.

The aim of this paper is to introduce a new notion of innovation for the analysis of nonlinear dynamics, and propose a representation of a time series as a function of current and lagged nonlinear innovations. The paper also presents a battery of specification tests based on corresponding nonlinear residuals for diagnostic checking of nonlinear models such as the NLARMA (nonlinear ARMA) models. A significant part of this work concerns the nonlinear impulse response analysis. The nonlinear innovations allow us to develop a new approach to study transitory and permanent shocks to models such as the popular GARCH or ACD. We also extend the application of impulse response analysis to the domain of financial strategies, and concider shock effects not only on the future values of the series of interest, but also on the outcome of a dynamic strategy, such as dynamic portfolio hedging, for example.

In section 2 we define a nonlinear gaussian innovation of a strongly stationary process and discuss the nonlinear regularity condition ensuring that the current and lagged values of the process contain information on current and lagged values of the innovation. Next, we derive a representation theorem for a nonlinearly regular stationary process where the current value of the process is expressed as a nonlinear function of current and lagged nonlinear innovations. In section 3 we consider parametric dynamic models and show how to find the nonlinear residuals by approximating nonlinear innovations. These residuals are next used to develop specification tests of the initial dynamic model, which extend the portmanteau tests introduced in the linear framework. Nonlinear ARMA models are defined in section 4, where we propose nonparametric estimation methods of nonlinear transforms of the autoregressive moving average. Section 5 is devoted to nonlinear impulse response analysis. We study the effects of permanent and transitory shocks to nonlinear gaussian innovations and compare our approach with impulse response techniques introduced by Gallant, Rossi, Tauchen (1993), and Koop, Pesaran, Potter (1996). Section 6 extends the impulse response techniques to a setup involving dynamic financial strategies. As an illustration we discuss the structural impulse response analysis for determining the Value at Risk and the minimum capital requirement under dynamic portfolio management.

2 Nonlinear gaussian innovation and representation theorem

In this section we consider an unidimensional strongly stationary process $(Y_t, t \in Z)$. We denote by $\mathcal{F}_t = \sigma(\underline{Y_t})$ the sigma algebra generated by the current and past values of the process. Moreover we assume:

Assumption A.1: The conditional distribution of Y_t given $\mathcal{F}_{t-1}$ is continuous on $[\mathcal{R}, \mathcal{B}(\mathcal{R})]$ with a positive p.d.f. denoted by f_{t-1}.

The associated c.d.f. F_{t-1} is continuous, strictly increasing and hence invertible.

2.1 Nonlinear gaussian innovations

Definition 1: The process $(\epsilon_t, t \in Z)$ is a nonlinear gaussian innovation of the process $(Y_t, t \in Z)$ if it satisfies the following conditions:

i) $(\epsilon_t, t \in Z)$ is a gaussian white noise IIN(0,1);

ii) ϵ_t and Y_t are in a continuous invertible relationship conditional on $\mathcal{F}_{t-1}$: $\epsilon_t = g_{t-1}(Y_t)$ a.s., where g_{t-1} is continuous, invertible and may depend on the past $\mathcal{F}_{t-1}$.

The second condition implies:

$$\sigma(\epsilon_t, \mathcal{F}_{t-1}) = \sigma(Y_t, \mathcal{F}_{t-1}), \quad \forall t, \tag{2.1}$$

and by recursion:

$$\sigma(\epsilon_t, \epsilon_{t-1}, ..., \epsilon_{t-p+1}, \mathcal{F}_{t-p}) = \sigma(Y_t, Y_{t-1}, ..., Y_{t-p+1}, \mathcal{F}_{t-p}), \quad \forall t, \ p \geq 0. \tag{2.2}$$

Property 2: Under assumption A.1, the strongly stationary process $(Y_t, t \in Z)$ admits a nonlinear gaussian innovation. It is unique up to a change of signs, date by date.

Proof:

i) The process $(\epsilon_t, t \in Z)$ defined by:

$$\epsilon_t = \Phi^{-1}[F_{t-1}(Y_t)], \quad t \in Z, \tag{2.3}$$

where Φ denotes the c.d.f. of the standard normal distribution, satisfies the two conditions of definition 1.

ii) Let us assume that ϵ_t^1 is another nonlinear gaussian innovation. Then there exists a continuous invertible relationship between ϵ_t and ϵ_t^1 conditional on: $\mathcal{F}_{t-1} : \epsilon_t = h_{t-1}(\epsilon_t^1)$, (say), and
$P[\epsilon_t^1 < \epsilon] = \Phi(\epsilon) = P[h_{t-1}(\epsilon_t^1) < \epsilon \mid \mathcal{F}_{t-1}]$.

The function h_{t-1} is continuous, invertible and therefore monotone. If it is increasing, conditional on $\mathcal{F}_{t-1}$, we get:

$$\Phi(\epsilon) = \Phi[h_{t-1}^{-1}(\epsilon)], \ \ \forall \epsilon,$$

which implies $h_{t-1} = Id$.

If it is decreasing, conditional on $\mathcal{F}_{t-1}$, we get:

$$\Phi(\epsilon) = 1 - \Phi[h_{t-1}^{-1}(\epsilon)], \ \ \forall \epsilon,$$

which implies $h_{t-1} = -Id$.

Q.E.D.

Therefore it is always possible to select the gaussian innovation process in order to obtain an increasing relationship between Y_t and ϵ_t at any time t. The corresponding innovation is uniquely defined by (2.3). Moreover this formula implies that the innovations of (Y_t) and of an invertible increasing transform of (Y_t) are identical.

2.2 Representation theorem

In this section we introduce a representation theorem which expresses the current value of the process as a function of current and lagged values of the innovation. By analogy to the Wold representation for linear processes, we first introduce the nonlinear regularity condition.

Definition 3: The $(Y_t, t \in Z)$ process is nonlinearly regular if $\mathcal{F}_{-\infty} = \cap_t \mathcal{F}_t$ is the degenerate sigma algebra.

This regularity condition implies in particular that [1]:

$$\lim_{h \to \infty} E(a(Y_t, Y_{t+1}, ..., Y_{t+q}) | \mathcal{F}_{t-h}) = Ea(Y_t, Y_{t+1}, ..., Y_{t+q}),$$

[1] It means that the process has short memory in the mean for any nonlinear transform using Granger's terminology [Granger (1995)].

for any integer q and integrable function a, which means that the initial value of the process is noninformative for a long horizon forecast. In the following property, the underlined variables denote processes up to and including the given date.

Property 4: If the strongly stationary process $(Y_t, t \in Z)$ satisfies assumption A.1 and is nonlinearly regular, then $\sigma(\underline{\epsilon_t}) = \sigma(\underline{Y_t}) = \mathcal{F}_t$.

Proof:

i) We have: $\sigma(\epsilon_t, ..., \epsilon_{t-p}) \subset \sigma(\underline{Y_t}),\ \ \forall p$, and then $\sigma(\underline{\epsilon_t}) = \vee_p\, \sigma(\epsilon_t, ..., \epsilon_{t-p}) \subset \sigma(\underline{Y_t})$.

ii) Conversely:

$$\sigma(\underline{Y_t}) = \sigma(\epsilon_t, ..., \epsilon_{t-p}) \vee \mathcal{F}_{t-p} \subset \sigma(\underline{\epsilon_t}) \vee \mathcal{F}_{t-p},\ \ \forall p.$$

Therefore $\sigma(\underline{Y_t}) \subset \cap_p \left[\sigma(\underline{\epsilon_t}) \vee \mathcal{F}_{t-p}\right] = \sigma(\underline{\epsilon_t}) \vee (\cap_p \mathcal{F}_{t-p}) = \sigma(\underline{\epsilon_t})$, due to the regularity condition.

Q.E.D.

The representation theorem is a consequence of the existence of simple hilbertian basis for gaussian processes. More precisely, let us introduce the Hermite polynomials:

$$H_j(\epsilon) = \sum_{0 \leq m \leq [j/2]} \frac{j!}{(j-2m)!m!2^m}(-1)^m \epsilon^{j-2m} \quad j = 0, 1, ... \tag{2.4}$$

A hilbertian basis of $L^2(\sigma(\underline{\epsilon_t}))$ is given by:

$$\frac{1}{\sqrt{j_1!}} H_{j_1}(\epsilon_{t-h_1}) \frac{1}{\sqrt{j_2!}} H_{j_2}(\epsilon_{t-h_2}) ... \frac{1}{\sqrt{j_n!}} H_{j_n}(\epsilon_{t-h_n}), \tag{2.5}$$

$n, j_1, ..., j_n, h_1, ..., h_n$ varying with $h_1 \neq h_2 .... \neq h_n$.

The representation theorem follows directly.

Property 5: If the strongly stationary process satisfies assumption A.1, is nonlinearly regular and square integrable, we get:

$$Y_t = \lim_{\substack{N \to \infty \\ J \to \infty \\ H \to \infty}} \sum_{n=1}^{N} \sum_{j_1, ..., j_n = 1, ..., J} \sum_{\substack{h_1, ..., h_n = 0, ..., H \\ h_1 \neq ... \neq h_n}} a^{(N,J,H)}_{j_1, ..., j_n, h_1, ..., h_n} H_{j_1}(\epsilon_{t-h_1}) H_{j_2}(\epsilon_{t-h_2}) ... H_{j_n}(\epsilon_{t-h_n}),$$

where Y_t is the mean square limit.

This representation theorem is of Volterra type [see, Volterra (1930), (1959), Nisio (1960), Priestley (1988)] and presents Y_t as a limit of polynomials in current and lagged values of a gaussian white noise. However our approach is closer to the lines followed by Wiener (1958), using Hilbert arguments.

The representation has an especially simple form, when the coefficients $a^{(N,J,H)}_{j_1,\dots,j_n,h_1,\dots,h_n}$ are independent of N, J, H. Indeed we get:

$$Y_t = \sum_{n=1}^{\infty} \sum_{j_1,\dots j_n=0}^{\infty} \sum_{h_1,\dots,h_n=0}^{\infty} a_{j_1,\dots,j_n,h_1,\dots,h_n} H_{j_1}(\epsilon_{t-h_1}) H_{j_2}(\epsilon_{t-h_2}) \dots H_{j_n}(\epsilon_{t-h_n}). \tag{2.6}$$

It is known that linear gaussian ARMA models satisfy this condition whenever the moving average part does not admit a root with unitary modulus [Whittle (1963)].

Finally note that the condition of square integrability of Y_t is not very restrictive. Indeed if Y_t is not square integrable we may find an increasing transformation $h(Y_t)$, which will satisfy this requirement. Next, the representation theorem can be applied to the process $h(Y_t)$ and by inverting h, a representation for Y_t will be obtained [where Y_t becomes now the limit in probability].

2.3 Example

The traditional nonlinear dynamic models introduced for financial applications such as ARCH models [Engle (1982)] or ACD models [Engle, Russell (1998)] usually contain nonlinear innovations in their specifications. As an illustration let us consider the ACD(1,1) model. The process of interest is a sequence of durations $\{Y_t,\, t \in \mathcal{Z}\}$. Let us introduce the conditional expectation of Y_t given the past: $\Psi_t = E(Y_t | \mathcal{F}_{t-1})$. It is assumed that the standardized durations Y_t/Ψ_t are independent with identical distributions whose c.d.f is F (say), and that Ψ_t satisfies the recursive equation:

$$\Psi_t = c + \alpha Y_{t-1} + \beta \Psi_{t-1}. \tag{2.7}$$

The nonlinear gaussian innovation is:

$$\epsilon_t = \Phi^{-1} F(Y_t/\Psi_t). \tag{2.8}$$

Therefore we can write:

$$Y_t = \Psi_t g(\epsilon_t), \tag{2.9}$$

where $g = F^{-1} \cdot \Phi$, and by substituting into the recursive equation (2.7) we get:

$$\begin{aligned} Y_t &= c g(\epsilon_t) + \alpha Y_{t-1} g(\epsilon_t) + \beta Y_{t-1} \frac{g(\epsilon_t)}{g(\epsilon_{t-1})} \\ &= \left[\alpha g(\epsilon_t) + \beta \frac{g(\epsilon_t)}{g(\epsilon_{t-1})}\right] Y_{t-1} + c g(\epsilon_t), \end{aligned} \tag{2.10}$$

which is an autoregressive representation with time dependent random autoregressive coefficients.

We may also write the recursive equation (2.7) as:

$$\Psi_t = c + \alpha \Psi_{t-1} g(\epsilon_{t-1}) + \beta \Psi_{t-1}, \tag{2.11}$$

which shows that the process (ϵ_t) is also the nonlinear innovation of the expectation process (Ψ_{t+1}).

3 Residual based diagnostics

In this section we consider a parametric model and define its nonlinear residuals. These residuals are next used to construct various specification tests.

3.1 Residuals

Let us consider a parametric model of the process $(Y_t, t \in Z)$, with the conditional p.d.f. parametrized by θ and denoted $f_{t-1}(.;\theta)$. The parameter can be estimated by the maximum likelihood, where the M.L. estimator is defined by:

$$\hat{\theta}_T = \underset{\theta}{Argmax} \sum_{t=1}^{T} \log f_{t-1}(Y_t;\theta). \tag{3.1}$$

We assume that standard regularity conditions are satisfied to ensure that this estimator is consistent, asymptotically normal and admits the asymptotic expansion:

$$\sqrt{T}(\hat{\theta}_T - \theta) = J^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \frac{\partial \log f_{t-1}}{\partial \theta}(Y_t;\theta) + o(\frac{1}{\sqrt{T}}), \tag{3.2}$$

where $J = E\left[-\frac{\partial^2 \log f_{t-1}(Y_t;\theta)}{\partial \theta \partial \theta'}\right]$.

The residuals of the model are defined by:

$$\hat{\epsilon}_t = \Phi^{-1}[F_{t-1}(Y_t;\hat{\theta}_T)]. \tag{3.3}$$

The specification tests are designed to verify whether these residuals satisfy moment conditions given below.

3.2 The moment conditions

We know that the model is well specified and θ_0 is the true parameter value, if and only if the variables $\epsilon_t = \Phi^{-1}[F_{t-1}(Y_t;\theta_0)]$ are IIN(0,1). This condition of gaussian white noise can be written in terms of moments of Hermite polynomials.

Property 6: $(\epsilon_t, t \in Z)$ is a gaussian white noise if and only if :

i) $EH_j(\epsilon_t) = 0, \;\; j \geq 1;$

ii) $EH_j^2(\epsilon_t) = 1, \;\; j \geq 1;$

iii) $E[H_j(\epsilon_t)H_k(\epsilon_t)] = 0, \;\; j,k \geq 1, \; j \neq k;$

iv) $E[H_j(\epsilon_t)\prod_{i=1}^{n} H_{k_i}(\epsilon_{t-h_i})] = 0, \;\; \forall n, \; k_1, ..., k_n, \; \forall h_1 \neq h_2 \neq ... \neq h_n.$

Proof:

Since the square integrable functions of $\underline{\epsilon_{t-1}}$ may be expanded in terms of a product of Hermite polynomials, even if the process (ϵ_t) is not gaussian, condition (iv) implies that:

$$\forall n, \; \forall f, g, : \;\; E[f(\epsilon_t)\, g(\epsilon_{t-1}, ..., \epsilon_{t-n})] = 0,$$

which is equivalent to the independence of ϵ_t with $\sigma(\underline{\epsilon_{t-1}})$.

Finally conditions i), ii), iii) imply that the marginal distribution of ϵ_t is standard normal.

Q.E.D.

3.3 Specification tests

We now introduce tests statistics based on the moment conditions outlined in the previous section. Let us denote

$$\xi_{j,k,h} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \{H_j(\hat{\epsilon}_t)H_k(\hat{\epsilon}_{t-h}) - \delta_{0,h}\delta_{j,k}\}, \;\; j,k \geq 0, \; h \geq 0, \tag{3.4}$$

where δ is the Kronecker symbol. Under the null hypothesis of correct specification, we obtain the following expansion of the test statistics [see, Appendix 1]:

$$\xi_{j,k,h} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} Z_t(j,k,h), \tag{3.5}$$

where :

$$Z_t(j,k,h) = H_j(\epsilon_t)H_k(\epsilon_{t-h}) - \delta_{0,h}\delta_{j,k} + A_{j,k,h}J^{-1}\frac{\partial \log f_{t-1}}{\partial \theta}(Y_t;\theta_0), \tag{3.6}$$

$$A_{j,k,h} = E[H_j(\epsilon_t)\tfrac{\partial H_k}{\partial \epsilon}(\epsilon_{t-h})\tfrac{1}{\phi(\epsilon_{t-h})}\tfrac{\partial F_{t-h-1}}{\partial \theta'}(Y_{t-h};\theta_0)] + E[\tfrac{\partial H_j(\epsilon_t)}{\partial \epsilon}H_k(\epsilon_{t-h})\tfrac{1}{\phi(\epsilon_t)}\tfrac{\partial F_{t-1}}{\partial \theta'}(Y_t;\theta_0)],$$

where ϕ denotes the p.d.f. of the standard normal distribution.

We note that if the model is well specified, $H_j(\epsilon_t)H_k(\epsilon_{t-h}) - \delta_{0,h}\delta_{j,k}$ and $\frac{\partial \log f_{t-1}(Y_t;\theta_0)}{\partial \theta}$ are martingale difference sequences. Therefore the components of $Z_t(j,k,h)$ are uncorrelated.

Property 7: If the parametric dynamic model is well specified the test statistics are asymptotically normal with zero mean and covariances given by:

$$Cov_{asy}[\xi_{j,k,h}, \xi_{j^*,k^*,h^*}] = Cov[Z_t(j,k,h), Z_t(j^*,k^*,h^*)].$$

In practice the various moment conditions can be considered sequentially. For instance

· for $j = k = 1, h$ varying, $h \geq 1$, the testing procedures will be based on the standard residual autocovariances;

· for $j = k = 2, h$ varying, $h \geq 1$, the testing procedures will be based on the sample autocorrelations of the squared residuals [see McLeod, Li (1983)];

· for $j = 2, k = 1, h$ varying, $h \geq 1$, the testing procedures allow to detect some remaining correlation between the squared residuals and the lagged residuals [which extends the suggestion by Lawrance, Lewis (1985)]....

However it has to be emphasized that even if the $\xi_{j,k,h}$ statistics are asymptotically independent, it is in general not possible to standardize them and obtain the usual form of the Liung-Box statistics. The reason is that we have to take into account the nonlinear dynamics which entails kurtosis and skewness effects and requires appropriate modifications of the asymptotic variances.

4 Nonlinear ARMA Models

Nonlinear ARMA models are introduced by distinguishing nonlinear effects of the lagged variables and nonlinear effects of the nonlinear innovations.

4.1 Definition

Definition 7: The strongly stationary process $(Y_t, t \in Z)$ has a nonlinear ARMA(p,q) or NLARMA(p,q) representation if and only if it satisfies a recursive relation [see Granger, Terasvirta (1993)]:

$$Y_t = g(Y_{t-1}, ..., Y_{t-p}, \epsilon_t, ..., \epsilon_{t-q}), \tag{4.1}$$

where $(\epsilon_t, t \in Z)$ is a gaussian nonlinear innovation, g is a function which is invertible with respect to ϵ_t, and is not constant with respect to Y_{t-p}, ϵ_{t-q} respectively. The coefficients p and q are the (nonlinear) autoregressive and moving average orders, respectively.

For $q = 0$, we get a nonlinear autoregression (NLAR) of order p [see Tong (1990), p 96]:

$$Y_t = g(Y_{t-1}, ..., Y_{t-p}, \epsilon_t) \Leftrightarrow c(Y_t, Y_{t-1}, ..., Y_{t-p}) = \epsilon_t, \quad \text{say}.$$

For $p = 0$ we get a nonlinear moving average (NLMA) of order q [see Tong (1990), p 115]:

$$Y_t = g(\epsilon_t, \epsilon_{t-1} ..., \epsilon_{t-q}).$$

In general, NLARMA models may contain cross effects of lagged values of the process and lagged innovations, as shown in the example of the ACD(1,1) model, which admits the NLARMA(1,1) representation (2.10). The presence of these cross effects makes it difficult to characterize the NLARMA representations simply in terms of their distributional properties as it is done in the linear framework where a process is said to have a linear MA(q) [resp AR(p)] representation if and only if the autocorrelations [resp. partial autocorrelations] cut off at lag $q+1$ [resp $p+1$]. We have instead the following characterization of nonlinear autoregressive processes.

Property 8 : A strongly stationary process admits a nonlinear autoregressive representation of order p if and only if it is Markov of order p.

Proof: The necessary condition is obvious. The sufficient condition results from the definition of the nonlinear gaussian innovation: $\epsilon_t = \Phi^{-1}[F_{t-1}(Y_t)]$, and the fact that F_{t-1} depends on the past through $Y_{t-1}, ..., Y_{t-p}$ only.

Q.E.D.

The property below details the link between the existence of a nonlinear MA representation and the zero correlations between nonlinear transformations of current and lagged observations whenever the lag is large enough.

Property 9: If the strongly stationary process admits a nonlinear moving average representation of order q, then the sigma algebras $\sigma(Y_t)$ and $\mathcal{F}_{t-q-1} = \sigma(\underline{Y_{t-q-1}})$ are independent.

However the condition of Property 9 is not sufficient due to cross effects as explained in the example below.

Example 4.1: Let us consider the NLAR(1) process whose conditional p.d.f. is:

$$f(y_t|\, y_{t-1}) = \frac{1}{2} 1_{[-1,1]}(y_t) \left[1 + \lambda \frac{3\sqrt{15}}{2} y_t \left(y_{t-1}^2 - \frac{1}{3} \right) \right],$$

with $|\lambda| < \frac{1}{\sqrt{15}}$. The marginal distribution of Y_t is the uniform distribution on the interval $[-1,\ 1]$. It is easy to show that the inequality $\lambda < \frac{1}{\sqrt{15}}$ ensures that $f(y_t|\, y_{t-1})$ is non-negative on the set of admissible values of y_t, y_{t-1}. Now let us consider the conditional distribution at horizon 2. We get:

$$\begin{aligned} f(y_t|\, y_{t-2}) &= \int_{-1}^{1} f(y_t|\, y_{t-1}) f(y_{t-1}|\, y_{t-2}) dy_{t-1} \\ &= \frac{1}{2} 1_{[-1,1]}(y_t) \int_{-1}^{1} \left[1 + \lambda \frac{3\sqrt{15}}{2} y_t \left(y_{t-1}^2 - \frac{1}{3}\right)\right] \left[1 + \lambda \frac{3\sqrt{15}}{2} y_{t-1} \left(y_{t-2}^2 - \frac{1}{3}\right)\right] dy_{t-1} \\ &= \frac{1}{2} 1_{[-1,1]}(y_t) = f_0(y_t). \end{aligned}$$

Therefore $\sigma(Y_t)$ is independent of $\sigma(\underline{Y_{t-2}})$. The AR function can be explicited by considering the conditional c.d.f.

$$F(y_t|\, y_{t-1}) = \frac{1}{2}(y_t + 1) + \lambda \frac{3\sqrt{15}}{8}(y_t^2 - 1)(y_{t-1}^2 - \frac{1}{3}),$$

and the equality

$$F(y_t|\, y_{t-1}) = \Phi(\epsilon_t) \Leftrightarrow y_t = \Psi(y_{t-1}, \epsilon_t).$$

It is easy to verify that

$$y_t = \Psi[\Psi(y_{t-2}, \epsilon_{t-1}), \epsilon_t],$$

actually depends on y_{t-2} and therefore the process does not admit a NLMA(1) representation.

4.2 Nonparametric estimation of nonlinear ARMA models

In this section we assume that $(Y_t,\, t \in Z)$ has a nonlinear ARMA(p,q) representation, and discuss nonparametric estimation of the function g for given orders p and q.

i) **Autoregressive processes**.

Let us first consider a NLAR(p) process. By applying appropriate nonlinear transformations, the models are:

$$c(Y_t, ..., Y_{t-p}) = \epsilon_t, \tag{4.2}$$

where $(\epsilon_t,\, t \in Z)$ is a gaussian nonlinear innovation. By construction, we know that:

$$c(Y_t, ..., Y_{t-p}) = \Phi^{-1}[F_{t-1}(Y_t)]. \tag{4.3}$$

Therefore a nonparametric estimator of the autoregressive function c is immediately deduced from a nonparametric estimator of the conditional c.d.f. F_{t-1}. For instance, we can consider a kernel estimator of the conditional c.d.f.:

$$\begin{aligned} \hat{F}_{t-1}(y_t) &= \hat{F}(y_t | Y_{t-1} = y_{t-1}, ..., Y_{t-p} = y_{t-p}) \\ &= \frac{\sum_{t=1}^{T} \left\{ 1_{Y_t < y_t} \prod_{j=1}^{p} \frac{1}{h} K \left[\frac{Y_{t-j} - y_{t-j}}{h} \right] \right\}}{\sum_{t=1}^{T} \prod_{j=1}^{p} \frac{1}{h} K \left[\frac{Y_{t-j} - y_{t-j}}{h} \right]}, \end{aligned} \tag{4.4}$$

where K is a second order kernel, i.e. satisfies $\int uK(u)du = 0$, $\int u^2 K(u)du < +\infty$.

Then we find that:

$$\hat{c}(y_t, \dots, y_{t-p}) = \Phi^{-1}[\hat{F}_{t-1}(y_t)]. \tag{4.5}$$

Obviously, these nonparametric techniques can only be implemented if the number of observations is large, compared to the autoregressive order p.

Under standard regularity conditions [see, e.g. Bosq (1998)] including the convergence of the bandwidth to zero at an appropriate rate, this functional estimator is consistent and asymptotically normal:

$$\sqrt{Th^p}[\hat{c}(y_t, ..., y_{t-p}) - c(y_t, ..., y_{t-p})] \to N\left[0, \frac{[\int u^2 K(u)du]^p}{\{\Phi[c(y_t, ..., y_{t-p})]\}^2} \frac{F_{t-1}(y_t)[1 - F_{t-1}(y_t)]}{f(y_{t-1}, ..., y_{t-p})}\right], \tag{4.6}$$

where $f(y_{t-1}, ..., y_{t-p})$ is the joint p.d.f. of $y_{t-1}, ..., y_{t-p}$

ii) **Approximation of ARMA processes by long autoregressive representations**

In the general case of ARMA processes of small orders p and q, we can follow the approach outlined below: [see, Gourieroux, Monfort (1997), p.188 for the analogue in the linear framework]

Step 1: We estimate an approximated autoregressive representation with a sufficiently large autoregressive order P:

$$C(Y_{t-1}, ..., Y_{t-P}) = \epsilon_t, \quad \text{(say)}.$$

We denote by $\hat{C}$ the associated estimator.

Step 2: We find the residuals of the model

$$\hat{\epsilon}_t = \hat{C}(Y_{t-1}, ..., Y_{t-P}), \quad t = 1, ..., T,$$

which are proxies for the gaussian innovations.

Step 3: We finally consider the nonlinear ARMA representation:

$$Y_t = g(Y_{t-1}, ..., Y_{t-p}, \epsilon_t, ..., \epsilon_{t-q}), \quad \text{(say)}.$$

The g function can be estimated by the regressogram of Y_t on $Y_{t-1}, ..., Y_{t-p}, \hat{\epsilon}_t, ..., \hat{\epsilon}_{t-q}$:

$$\hat{g}(y_{t-1}, ..., y_{t-p}, \epsilon_t, ..., \epsilon_{t-q}) \tag{4.7}$$

$$= \frac{\sum_{t=1}^{T} \left\{ Y_t \prod_{i=1}^{p} \frac{1}{h} K\left[\frac{Y_{t-i} - y_{t-i}}{h}\right] \prod_{j=1}^{q} \frac{1}{h} K\left[\frac{\hat{\epsilon}_{t-j} - \epsilon_{t-j}}{h}\right] \right\}}{\sum_{t=1}^{T} \frac{1}{h} K\left[\frac{Y_{t-i} - y_{t-i}}{h}\right] \prod_{j=1}^{q} \frac{1}{h} K\left[\frac{\hat{\epsilon}_{t-j} - \epsilon_{t-j}}{h}\right]}, \tag{4.8}$$

since, in particular, the regressogram can be applied to estimate a deterministic relationship [Bosq, Guegan (1995)].

5 Impulse Response Analysis

5.1 Background

In recent literature, Gallant, Rossi, Tauchen (1993), and Koop, Pesaran, Potter (1996), [henceforth GRT and KPP] have proposed extensions of the traditional impulse response analysis to nonlinear dynamic models. Both papers emphasize the specificity of nonlinear framework for impulse response functions, considered as the time profile of the shock effect on the behaviour of the series [see, e.g. KPP (1996)].

i) For ARIMA models, impulse responses have a symmetry property (i.e. a transitory shock of $-\delta$ has exactly the opposite effect of a transitory shock of $+\delta$) , whereas in the nonlinear case the effects of opposite shocks may be very different.

ii) In the linear framework we have the property of "shock linearity" [i.e. a transitory shock of $k\delta$ has k times the effect of a shock of δ], whereas the effect of the magnitude of the shock is nonlinear in the general case.

iii) For ARIMA models the effect of the shocks does not depend on the past history, whereas this path dependency is crucial in nonlinear framework.

iv) Finally, it is necessary to consider some distributional properties of the impulse response function, and not only the effect of the shocks on the conditional expectation of the future variable of interest.

However while both extensions consider only transitory shocks, their proposed impulse response analysis differ.

i) The GRT analysis is performed conditional on the observed history, whereas KPP propose to integrate out the possible histories.

ii) Fundamental differences arise from definitions of the transitory shocks. In the spirit of the Keynesian multiplier analysis, GRT propose to shock directly the variable, and not an innovation. The drawback of this approach is twofold. It can not be extended to an analysis of permanent shocks, and the idea of symmetric shocks is unclear. Indeed symmetric shocks to innovations do not necessarily correspond to symmetric shocks to the current variable. On the contrary, KPP following Sims (1978) [see also Blanchard, Quah (1989)], consider shocks to innovations. However they consider innovations defined by conditionally centering and rescaling the variable Y_t; these innovations:

$$v_t = (V_{t-1}Y_t)^{-\frac{1}{2}}(Y_t - E_{t-1}Y_t),$$

are not sufficiently corrected for the presence of nonlinear temporal dependence. Firstly these innovations may feature temporal dependence in moments of order larger then three, which may affect the interpretation of impulse responses. Secondly, the conditional distribution of v_t is not symmetric in general, and this could affect the interpretation of symmetric shocks.

The difficulties encountered by these authors are due to ambiguous definitions of innovations in nonlinear framework. Using the gaussian innovations which have been uniquely defined in section 2 , we can now propose a complete innovation based impulse response analysis for both transitory and permanent shocks.

5.2 Definitions of the impulse response functions

The impulse response analysis can be based on the Volterra type decomposition (see, property 5), where:

$$Y_t = a_t(\epsilon_t, \epsilon_{t-1}, ..., \epsilon_1, \underline{\epsilon_0}), \tag{5.1}$$

and (ϵ_t) is a gaussian white noise, with unitary variance. Since the distribution of ϵ_t is symmetric, the shocks of δ and $-\delta$ have the same infinitesimal occurrence. As well, since the distribution of ϵ_t is independent of time, the shocks of the same magnitude δ at different dates have also the same infinitesimal occurrence, which allows to consider "permanent" shocks.

As suggested by GRT (1993) the analysis needs to be conditioned on the history before the shocks. Therefore if the shocks hit the process at date 1, the previous values of the process and the innovations are known, i.e. $\underline{\epsilon_0}$ is fixed. Then, at date 0, we have to evaluate the effect of a sequence of deterministic shocks $\delta_1, \delta_2, ..., \delta_t, ...$ occurring at future dates on the future profile of the process. These effects have to be measured with respect to a benchmark which is the path followed under the absence of shocks. Since future innovations are unknown, this benchmark is random. We denote by: $\epsilon_1^s, \epsilon_2^s, ..., \epsilon_t^s, ...$ a future path for the innovations, where $\epsilon_1^s, \epsilon_2^s, ..., \epsilon_t^s, ...$ are IIN (0,1) conditional on $\underline{\epsilon_0}$. The random benchmark is:

$$Y_t^s(\underline{\epsilon_0}) = a_t(\epsilon_t^s, \epsilon_{t-1}^s, ..., \epsilon_1^s, \underline{\epsilon_0}), \tag{5.2}$$

whereas the profile after shocks arrival is:

$$Y_t^s(\underline{\delta}, \underline{\epsilon_0}) = a_t(\epsilon_t^s + \delta_t, \epsilon_{t-1}^s + \delta_{t-1}, ..., \epsilon_1^s + \delta_1, \underline{\epsilon_0}), \tag{5.3}$$

where $\underline{\delta} = (\delta_1, ..., \delta_t, ...)$.

The entire effect of the sequence of shocks is summarized by the joint path distribution of:

$$[Y_t^s(\underline{\epsilon_0}), , Y_t^s(\underline{\delta}, \underline{\epsilon_0}),\ t \geq 1].$$

In practice we have to select sequences of shocks and summary statistics of the joint path distribution. The standard response analysis concerns:

either transitory shocks at date 1: $\delta_1 = \delta, \delta_t = 0, t \geq 2$,

or transitory shocks at date t_0: $\delta_{t_0} = \delta,\ \delta_t = 0,\ t \neq t_0$,

or permanent shock starting at date 1: $\delta_t = \delta,\ \ t \geq 2$.

They differ in terms of the sign and magnitude of δ.

The standard distributional summary statistics considered in the literature [GRT (1993), KPP(1996)] are:

either differences of expectations of the series:

$$E[Y_t^s(\underline{\delta}, \underline{\epsilon_0})|\underline{\epsilon_0}] - E[Y_t^s(\underline{\epsilon_0})|\underline{\epsilon_0}],$$

or differences of expectations of transformed series:

$$E[g(Y_t^s(\underline{\delta},\underline{\epsilon_0}))|\underline{\epsilon_0}] - E[g(Y_t^s(\underline{\epsilon_0}))|\underline{\epsilon_0}],$$

where g is a given nonlinear function,

or differences of variances:

$$V[Y_t^s(\underline{\delta},\underline{\epsilon_0})|\underline{\epsilon_0}] - V[Y_t^s(\underline{\epsilon_0})|\underline{\epsilon_0}].$$

5.3 Nonlinear AR(1) process

i) **The dynamics**.

As an illustration we consider a dynamic NLAR(1) model defined by:

$$Y_t = g(Y_{t-1};\epsilon_t) = g^{*(1)}(Y_{t-1};\,\epsilon_t), \quad \text{say}, \tag{5.4}$$

where (ϵ_t) is a standard gaussian white noise and gn an invertible function with respect to ϵ. By recursion we find the expression of Y_t as a function of Y_0 and innovations $\epsilon_1^t = (\epsilon_1, ..., \epsilon_t)$:

$$Y_t = g^{*(t)}[Y_0, \epsilon_1^t], \tag{5.5}$$

where $g^{*(t)}$ is recursively defined by:

$$\begin{aligned} g^{*(t)}[Y_0,\epsilon_1^t] &= g\{g^{*(t-1)}[Y_0,\epsilon_1^{t-1}],\epsilon_t\} &(5.6)\\ &= g^{*(t-1)}\{g[Y_0,\epsilon_1],\epsilon_2^t\}. &(5.7)\end{aligned}$$

The equality (5.6) implies:

$$\frac{\partial g^{*(t)}}{\partial y}[Y_0,\epsilon_1^t] = \frac{\partial g}{\partial y}[Y_{t-1},\epsilon^t]\frac{\partial g^{*(t-1)}}{\partial y}[Y_0,\epsilon_1^{t-1}],$$

and by the chain rule:

$$\frac{\partial g^{*(t)}}{\partial y}[Y_0,\epsilon_1^t] = \prod_{\tau=1}^{t}\frac{\partial g}{\partial y}[Y_{\tau-1},\epsilon_\tau]. \tag{5.8}$$

ii) **Local impact of a transitory shock at date 1**.

Let us consider a small transitory shock $\delta_1 = \delta$. We get:

$$\begin{aligned} Y_t(\delta) &= g^{*(t)}[Y_0,\epsilon_1^t + \delta_1^t]\\ &= g^{*(t-1)}[g(Y_0,\epsilon_1+\delta),\epsilon_2^t], \quad \text{from (5.7)},\end{aligned}$$

$$\approx \quad g^{*(t-1)}[Y_1, \epsilon_2^t] + \frac{\partial g^{*(t-1)}}{\partial y}[Y_1, \epsilon_2^t]\frac{\partial g}{\partial y}[Y_{\tau-1}, \epsilon_\tau]\delta$$
$$= \quad Y_t + \prod_{\tau=1}^{t} \frac{\partial g}{\partial y}[Y_{\tau-1}, \epsilon_\tau]\delta, \quad \text{from (5.8)}.$$

If we consider a nonlinear transformation of the process, H say, we get:

$$H[Y_t(\delta)] \approx H(Y_t) + \frac{dH}{dy}(Y_t)\prod_{\tau=1}^{t}\frac{\partial g}{\partial y}[Y_{\tau-1}, \epsilon_\tau]\delta. \tag{5.9}$$

Therefore the infinitesimal effect of the transitory shock on the expectation of $H(Y_t)$ conditional on Y_0 is:

$$\frac{1}{\delta}E_0\{H[Y_t(\delta)] - H(Y_t)\} = E_0\left[\frac{dH}{dy}(Y_t)\prod_{\tau=1}^{t}\frac{\partial g}{\partial y}[Y_{\tau-1}, \epsilon_\tau]\right]. \tag{5.10}$$

It is interesting to consider the long run impact of a transitory shock, i.e. the behaviour of either $\frac{1}{\delta}[Y_t(\delta) - Y_t]$, or $\frac{1}{\delta}E_0[Y_t(\delta) - Y_t]$, when t tends to infinity. Indeed Nelson (1990) [see also Bougerol, Picard (1992)] has shown in the framework of GARCH models, that the responses to shocks can be significantly different. More precisely, if the process is nonlinearly regular, the asymptotic impact of a shock on Y_0 is equal to zero. For large t this impact is:

$$\prod_{\tau=1}^{t}\frac{\partial g}{\partial y}[Y_{\tau-1}, \epsilon_\tau]$$
$$= \quad \exp\left\{\sum_{\tau=1}^{t}\left[\frac{\log \partial g}{\partial y}(Y_{t-1}, \epsilon_\tau)\right]\right\}$$
$$\sim \quad \exp\left[tE\log\frac{\partial g}{\partial y}(Y_{\tau-1}, \epsilon_\tau)\right],$$

and tends to zero if $E\log\frac{\partial g}{\partial y}(Y_{\tau-1}, \epsilon_\tau) < 0$, which is a necessary condition for a strongly stationary regular process. $E\log\frac{\partial g}{\partial y}(Y_{\tau-1}, \epsilon_\tau)$ is the Liapunov exponent of the dynamic system [Oseledec (1968)].

However, by taking the expectation and using the convexity inequality, we get:

$$E\left[\prod_{\tau=1}^{t}\frac{\partial g}{\partial y}(Y_{\tau-1}, \epsilon_\tau)\right]$$
$$= \quad E\left[\exp\sum_{\tau=1}^{t}\log\frac{\partial g}{\partial y}(Y_{\tau-1}, \epsilon_\tau)\right]$$
$$\geq \quad \exp\left[E\sum_{\tau=1}^{t}\log\frac{\partial g}{\partial y}(Y_{\tau-1}, \epsilon_\tau)\right]$$

$$\approx \quad \exp\left[t\, E \log \frac{\partial g}{\partial y}(Y_{\tau-1}, \epsilon_\tau)\right],$$

and the stationarity condition $E\left[\log \frac{\partial g}{\partial y}(Y_{\tau-1}, \epsilon_\tau)\right] < 0$ does not necessarily imply that the impulse response vanishes in average for large t.

iii) **Linear AR(1) model with a random autoregressive coefficient**.

Let us consider the dynamic bilinear model [Tong (1993), p.8]:

$$Y_t = (a + b\epsilon_t)Y_{t-1} + \epsilon_t, \quad t \text{ varying},$$

where (ϵ_t) is a standard gaussian white noise. We can explicitely compute the effect of a transitory shock hitting at date 1. Indeed the disturbed path is such that:

$$Y_t^D = (a + b\epsilon_t)Y_{t-1}^D + \epsilon_t, \quad \forall t \geq 2.$$

We see that:

$$\begin{aligned} \Delta Y_t = Y_t^D - Y_t &= (a + b\epsilon_t)\Delta Y_{t-1} \\ &= \prod_{\tau=2}^{t}(a + b\epsilon_\tau)\Delta Y_1 \\ &= \prod_{\tau=2}^{t}(a + b\epsilon_\tau)(1 + bY_0)(\delta\epsilon_1). \end{aligned}$$

This model satisfies the property of shock linearity because the effect of the shock is a linear function of $\delta\epsilon_1$. Moreover we know from Bougerol, Picard (1992) that, for large t, the coefficient:

$$\begin{aligned} &\prod_{\tau=2}^{t}(a + b\epsilon_\tau)(1 + bY_0) \\ &= \exp\left\{(t-1)\frac{1}{t-1}\sum_{\tau=2}^{t}\log(a + b\epsilon_\tau)\right\}(1 + bY_0) \\ &\sim \exp[(t-1)E\log(a + b\epsilon_\tau)](1 + bY_0) \end{aligned}$$

tends to zero if and only if $E\log(a + b\epsilon_\tau) < 0$. However the effect of the shock on the expectation of Y_t is :

$$E[\Delta Y_t | Y_0] = a^{t-1}(1 + bY_0)\delta\epsilon_1.$$

This average effect tends to zero if $|a| < 1$, which is a more stringent condition than the negativity of $E \log(a + b\epsilon_\tau)$, due to the convexity inequality.

Example 5.1: We note that the ACD(1,1) model introduced in subsection (2.3) belongs to this class. Indeed, the sequence of expected durations satisfies:

$$\Psi_t = c + (\alpha g(\epsilon_{t-1}) + \beta)\Psi_{t-1},$$

and the impact of a transitory shock is such that :

$$\Delta\Psi_t = (\alpha g(\epsilon_{t-1}) + \beta)\Delta\Psi_{t-1}, \text{ for } t \geq 3.$$

This effect asymptotically vanishes in average if

$$|E(\alpha g(\epsilon) + \beta)| = |\alpha + \beta| < 1,$$

and vanishes path by path if the Liapunov exponent is negative $E \log(\alpha g(\epsilon)+\beta) < 0$. These conditions imply restrictions on α, β and on the pattern of the distribution of standardized durations. For instance, if $g(\epsilon)$ has an exponential distribution, the Liapunov exponent is

$$\int_0^\infty \log(\alpha x + \beta) \exp -x dx = \log \beta + \exp(\beta/\alpha) E_1(\beta/\alpha),$$

where $E_1(x) = \int_x^\infty \frac{\exp -t}{t} dt$ is the exponential integral [see Abramowitz, Stegun (1964), formula 5.1.1, page 228]. We plot in Figure 5.1 the set of all points (α, β) for which the Liapunov exponent is zero.

Insert Figure 5.1

Below this frontier, all pairs of (α, β) coordinates are associated to negative Liapunov exponents, while the coordinates (α, β) above it are associated to its positive values. We find that the frontier is decreasing from $(\alpha - 0.0, \beta = 1.0)$ down to $(\alpha \approx 1.4, \beta \approx 0.1)$. For β's greater than 1.4 the frontier approches asymptotically the axis of α.

5.4 Simulation results

In this section we illustrate the computation and analysis of impulse response functions using the examples of an ACD(1,1) model and a factor model with a distinct form of nonlinear temporal dependence.

i) **The autoregressive conditional duration (ACD) model.**

We consider the ACD(1,1) model of subsection (2.3) with an exponential distribution of the standardized duration. The initial values have been fixed to $\epsilon_0 = 0.0, y_0 = 2.0$. We perform two experiments involving two sets of parameter values:

· experiment 1: $c = 1,\ \alpha = 0.3,\ \beta = 0.2,$

· experiment 1: $c = 1,\ \alpha = 0.4,\ \beta = 0.64.$

In the first experiment the shock effect asymptotically vanishes in the mean and path by path, whereas in the second experiment we observe a different outcome. We consider transitory shocks δ occuring at date 1, and taking values $\delta = -1, -0.9, ...0.9, 1$, with the benchmark corresponding to $\delta = 0.0$. The maximal horizon is $H = 10$.

We display in Figure 5.2 the joint simulated paths for the benchmark and two perturbed series with $\delta = +/-1$, for both experiments. The effects of shocks quickly dissipate in experiment 1 whereas they are more persistent in experiment 2 even though they also vanish asymptotically.

Insert Figure 5.2: Simulated Paths.

We plot in Figure 5.3 the (marginal) distribution of $Y_t(\delta)$ for horizon $t = 3$, and $\delta = -1, 0, +1$, computed conditionally on the information available at date 0.

Insert Figure 5.3: Marginal Distribution at Horizon 3.

We note that the shocks have an effect on the means and tails of the distribution. These effects can be evaluated by considering two summary statistics:

- the mean deviation $EY_t(\delta) - EY_t$,

- the variance of the deviation from benchmark: $V(Y_t(\delta) - Y_t)$,

for different horizons $t = 1, ..., 10$ and different values of transitory shocks. They are shown in Figures 5.4 and 5.5 for experiment 1, and in Figures 5.6 and 5.7 for experiment 2.

Insert Figure 5.4: Mean Deviation from the Benchmark.

Insert Figure 5.5: Variance of the Deviation from Benchmark.

Insert Figure 5.6: Mean Deviation from the Benchmark.

Insert Figure 5.7: Variance of the Deviation from Benchmark.

We observe explosive patterns of the averaged effects of shocks in the second experiment, although we found earlier no explosive paths. The similar patterns of responses associated to different shocks are due to the simple formula of the deviation from the benchmark for the ACD(1,1) model. Indeed it is easy to see that :

$$Y_t(\delta) - Y_t = \alpha \prod_{\tau=3}^{t} \left[\alpha g(\epsilon_\tau) + \beta \frac{g(\epsilon_\tau)}{g(\epsilon_{\tau-1})}\right] g(\epsilon_2) \left[c + \left(\alpha + \frac{\beta}{g(\epsilon_0)}\right) y_0\right] [g(\epsilon_1 + \delta) - g(\epsilon_1)],$$

which implies :

$$E[Y_t(\delta) - Y_t] = A_t E[g(\epsilon_1 + \delta) - g(\epsilon_1)],$$

where A_t is a positive number depending on the horizon. For the same reason we get:

$$\begin{aligned} V[Y_t(\delta) - Y_t] &= E[Y_t(\delta) - Y_t]^2 - (E[Y_t(\delta) - Y_t])^2 \\ &= B_t E[g(\epsilon_1 + \delta) - g(\epsilon_1)]^2 - A_t^2 \left(E[g(\epsilon_1 + \delta) - g(\epsilon_1)]\right)^2 . \end{aligned}$$

The response function depends on both the magnitude of the shock and the horizon. To clearify the dependence with respect to the shock size we reproduce in Figures 5.8-5.9 the Figures 5.4-5.5, with δ measured on the x-axis. We find that the response function is convex for the mean deviation with a stronger convexity asssociated to negative shocks. In particular the properties of symmetry and linearity of linear impulse responses are not satisfied. The variance function displays an asymmetric effect of positive and negative shocks.

Insert Figure 5.8: Mean Deviation from the Benchmark.

Insert Figure 5.9: Variance of the Deviation from Benchmark.

An impulse response analysis based only on the differences $Y_t(\delta) - Y_t$ can be misleading since it does not represent the shock effect with respect to the mean of transformed series. It is more informative to consider the joint bivariate distribution of $Y_t(\delta), Y_t$ and examine how it depends on the magnitude of the shock and on the horizon. The corresponding scatterplots are given in Figure 5.10 for $\delta = +1, -1$, and horizon 3.

Insert Figure 5.10: Scatterplot at Horizon 3.

ii)**A model with linear autoregressive factors.**

We introduce a gaussian AR(1) model:

$$Z_t = \rho Z_{t-1} + \epsilon_t, \quad t \quad \text{varying}, \tag{5.11}$$

and the process of interest defined by:

$$Y_t = a(\epsilon_t, Z_t), \quad t \quad \text{varying}, \tag{5.12}$$

where a is a given function. This process is generated by two underlying factors ϵ_t, Z_t based on the same gaussian white noise. For a transitory shock δ hitting the process at date 1, we get:

$$Y_t(\delta) = a(\epsilon_t, Z_t + \rho^{t-1}\delta), \;\; t \geq 2,$$

and the joint distribution of $[Y_t, Y_t(\delta)]$ can easily be deduced from the joint gaussian distribution of ϵ_t, Z_t. The effect of the shock depends on the nonlinear transformation a.

Let us first consider the function:

$$a(\epsilon_t, Z_t) = sign(\epsilon_t) \exp Z_t. \tag{5.13}$$

We can see that: $Y_t(\delta) = Y_t \exp(\rho^{t-1}\delta)$, and the joint distribution of $[Y_t, Y_t(\delta)]$ is degenerate so that its support is a line passing through the origin. The shock has no effect on the sign of Y_t, whereas it has a multiplicative effect on its absolute value.

Less extreme examples corresponding to the functions: $a(\epsilon_t, Z_t) = \epsilon_t Z_t$ and $a(\epsilon_t, Z_t) = Z_t/\epsilon_t$, respectively are illustrated in Figures 5.11-5.11. The values of the parameters are: $\rho = 0.9, \;\; \delta = +1, -1$, and the horizon is $t = 4$.

Insert Figure 5.11 Scatterplot a = epsilon*Z

Insert Figure 5.12 Scatterplot a = Z / epsilon

6 Value at Risk for a dynamic financial strategy

The specification and estimation of dynamic models for economic or financial series of interest is often a preliminary step before decision making involving dynamic strategies and market interventions. In such a framework we are more interested in consequences of shocks for the outcomes of the dynamic strategy than in their effect on the underlying series. In this section, we discuss this structural interpretation of impulse response functions using the example of the Value at Risk [VaR] employed in finance to measure and control the risks associated to a portfolio. In the first subsection we recall the standard definition of the VaR, and extend this definition to a dynamic strategy of risk assessment in the second subsection.

6.1 Definition of the VaR

We consider at date T a portfolio including the quantities $a_{0,T}$ and a_T of a riskfree asset and various risky assets, respectively. We denote by $y_{0,T+h}, y_{T+h},\ h = 0, 1, ..., H$ the future values of the assets. The VaR is in practice defined for portfolios, whose allocations are constrained to be fixed in the future [the so-called crystallization of the portfolio]. These future portfolio values are:

$$W_{T+h} = a_{0,T} y_{0,T+h} + a'_T y_{T+h},\ \ h = 0, 1, ..., H. \tag{6.1}$$

These values are random conditional on the information available at time T.

The Value at Risk of this portfolio evaluated at time T, for the horizon H and the critical value $\alpha \in [0, 1]$ is the quantity VaR [T, H, α] defined by:

$$P_T[W_{T+H} > -VaR\ \ [T, H, \alpha]] = 1 - \alpha, \tag{6.2}$$

where P_T is the conditional distribution of future prices. The VaR may be used to determine the minimal capital requirement to ensure that the total wealth, including the portfolio value and capital requirement, remains positive with a sufficiently large probability. More precisely, let us assume that the capital requirement can be invested at a rate providing a zero coupon price $B[T, H]$ for horizon H. Consequently, the VaR determined capital requirement woul be:

$$R[T, H, \alpha] = \frac{VaR\ \ [T, H, \alpha]}{B[T, H]}. \tag{6.3}$$

Although the risk should be optimally measured ex ante, in practice it is usually measured ex post. The reason for it is the belief that the conditional distribution P_T can be well approximated by the historical distribution estimated from recent past data[see Morgan (1994)]. However when a dynamic model of the asset prices is available, ex-ante computation of the VaR can be performed by simulation. For instance let us suppose that prices follow a (nonlinear) AR(1) model:

$$y_t = g(y_{t-1}, \epsilon_t),\ \ t\ \text{ varying}, \tag{6.4}$$

where g is a given function, and $y_{0,T+h} = 1,\ \forall h = 1, ..., H$. We can simulate the future risky asset prices given the current value y_T by:

$$y^s_{T+h} = g(y^s_{T+h-1}, \epsilon^s_{T+h}),\ \ h = 1, ..., H,\ s = 1, ..., S, \tag{6.5}$$

where $y^s_T = y_T$, and we deduce the simulated future values of the portfolio by:

$$W^s_{T+h} = W_T + a'_T(y^s_{T+h} - y_T), \quad h = 1, ..., H, \; s = 1, ..., S. \tag{6.6}$$

Then the VaR can be computed from the empirical distribution of these simulated values $W^s_{T+h}, \; s = 1, ..., S$.

Let us now introduce a shock δ to the innovation at date $T + 1$. We can perform similar computations after replacing ϵ^s_{T+1} by $\epsilon^s_{T+1} + \delta$, deduce the distribution of the future portfolio values $W^s_{T+h}(\delta)$, say, and the associated Value at Risk: VaR [T, H, α; δ]. This will allow us to study the sensitivity of the VaR and of the capital requirement to a transitory shock δ.

6.2 Extension to dynamic strategies

The approach presented above, is the one proposed by the regulators, which is not optimal in practice. Indeed, it is based on the assumption of fixed future portfolio allocations, whereas the investors regularly update these allocations to take advantage of price movements. A typical example is a hedging portfolio for an european call. The portfolio is often updated at regular dates $T + h, \; h = 1, ..., H$ (say), with allocations determined by the deltas of the Black-Scholes formula. Even though the price evolution in the Black-Scholes model obeys a linear dynamics, the deltas are complicated nonlinear functions of the prices, and the major part of the risk is due to these nonlinear adjustments.

Let us still assume the price of the riskfree asset $y_{0,T+h} = 1, \; h = 1, .., H$. The future values of a self-financed portfolio are:

$$W_{T+H}[a(.)] = W_T + \sum_{h=1}^{H} a_h(\underline{y_{T+h-1}})[y_{T+h} - y_{T+h-1}], \tag{6.7}$$

where $a_h(\underline{y_{T+h-1}})$ are the allocations in the risky assets considered at the h^{th} updating, which are path dependent in general.

For a given dynamic strategy $a(.) = [a_h(.), \; h = 1, ..., H]$, we can compute, like in the previous subsection, the simulated future portfolio values

$$W^s_{T+H}[a(.)] = W_T + \sum_{h=1}^{H} a_h(\underline{y^s_{T+h-1}})[y^s_{T+h} - y^s_{T+h-1}], \tag{6.8}$$

and deduce the VaR under and without a transitory shock. These Values at Risk will depend on the selected dynamic strategy. Let us denote them by

$$VaR\ [T, H, \alpha;\ a(.)] \text{ and } VaR\ [T, H, \alpha;\ a(.); \delta].$$

They can be compared to the Values at Risk evaluated for a portfolio with fixed allocations $a_h(.) = a_T,\ \forall h$:

$$VaR\ [T, H, \alpha] \text{ and } VaR\ [T, H, \alpha; \delta],$$

using the notation of subsection 6.1.

Appendix 1

Asymptotic properties of the test statistics

1. Expansion:

We first consider an expansion of the residual:

$$\begin{aligned}
\hat{\epsilon}_t &= \Phi^{-1}[F_{t-1}(Y_t;\hat{\theta}_T)] \\
&= \Phi^{-1}[F_{t-1}(Y_t;\theta_0) + \frac{\partial F_{t-1}}{\partial \theta'}(Y_t;\theta_0)(\hat{\theta}_T - \theta_0) + o(\frac{1}{\sqrt{T}})] \\
&= \Phi^{-1}[F_{t-1}(Y_t;\theta_0)] + \frac{1}{\phi[\Phi^{-1}(F_{t-1}(Y_t;\theta_0))]}\frac{\partial F_{t-1}}{\partial \theta'}(Y_t;\theta_0)(\hat{\theta}_T - \theta_0) + o(\frac{1}{\sqrt{T}}) \\
&= \epsilon_t + \frac{1}{\phi(\epsilon_t)}\frac{\partial F_{t-1}}{\partial \theta'}(Y_t;\theta_0)(\hat{\theta}_T - \theta_0) + o(\frac{1}{\sqrt{T}}).
\end{aligned}$$

We deduce the expansion of the test statistic:

$$\begin{aligned}
\xi_{j,k,h} &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}[H_j(\hat{\epsilon}_t)H_k(\hat{\epsilon}_{t-h}) - \delta_{0,h}\delta_{j,k}] \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\{H_j\left[\epsilon_t + \frac{1}{\phi(\epsilon_t)}\frac{\partial F_{t-1}}{\partial \theta'}(Y_t;\theta_0)(\hat{\theta}_T - \theta_0) + o(\frac{1}{\sqrt{T}})\right] \\
&\quad H_k\left[\epsilon_{t-h} + \frac{1}{\phi(\epsilon_{t-h})}\frac{\partial F_{t-h-1}}{\partial \theta'}(Y_{t-h};\theta_0)(\hat{\theta}_T - \theta_0) + o(\frac{1}{\sqrt{T}})\right] - \delta_{0,h}\delta_{j,k}\} \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}[H_j(\epsilon_t)H_k(\epsilon_{t-h}) - \delta_{0,h}\delta_{j,k}] \\
&\quad + \frac{1}{\sqrt{T}}\sum_{t=1}^{T} H_j(\epsilon_t)\frac{\partial H_k}{\partial \epsilon}(\epsilon_{t-h})\frac{1}{\phi(\epsilon_{t-h})}\frac{\partial F_{t-h-1}}{\partial \theta'}(Y_{t-h};\theta_0)(\hat{\theta}_T - \theta_0) \\
&\quad + \frac{1}{\sqrt{T}}\sum_{t=1}^{T} \frac{\partial H_j}{\partial \epsilon}(\epsilon_t)H_k(\epsilon_{t-h})\frac{1}{\phi(\epsilon_t)}\frac{\partial F_{t-1}}{\partial \theta'}(Y_t;\theta_0)(\hat{\theta}_T - \theta_0) + o(1) \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}[H_j(\epsilon_t)H_k(\epsilon_{t-h}) - \delta_{0,h}\delta_{j,k}] + A_{j,k,h}\sqrt{T}(\hat{\theta}_T - \theta_0) + o(1),
\end{aligned}$$

where:

$$A_{j,k,h} = E[H_j(\epsilon_t)\frac{\partial H_k}{\partial \epsilon}(\epsilon_{t-h})\frac{1}{\phi(\epsilon_{t-h})}\frac{\partial F_{t-h-1}}{\partial \theta'}(Y_{t-h};\theta_0)] + E[\frac{\partial H_j(\epsilon_t)}{\partial \epsilon}H_k(\epsilon_{t-h})\frac{1}{\phi(\epsilon_t)}\frac{\partial F_{t-1}}{\partial \theta'}(Y_t;\theta_0)].$$

Figure 5.1: Liapunov Exponent = 0 Frontier

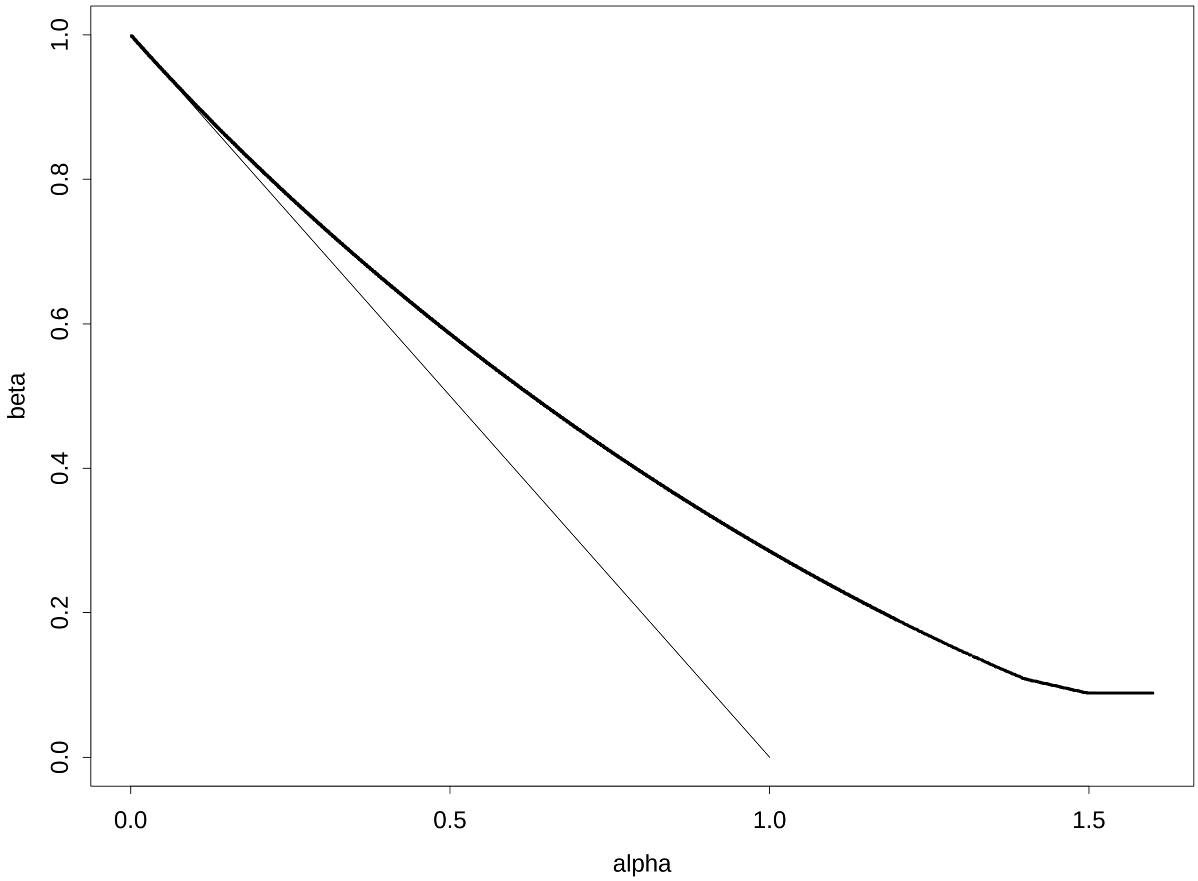


Figure 5.2: Simulated Paths

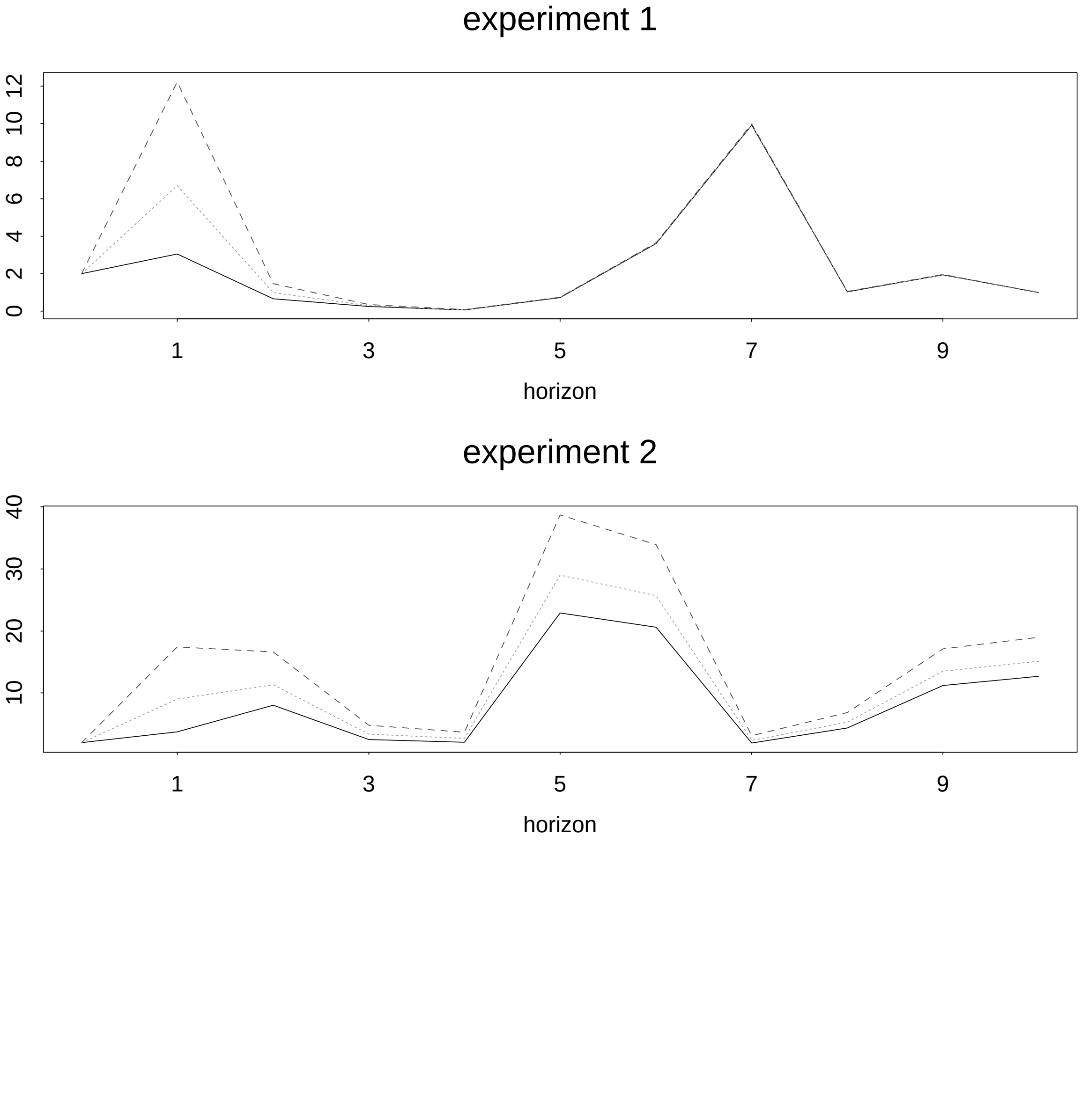


Figure 5.3: Marginal distributions at horizon 3

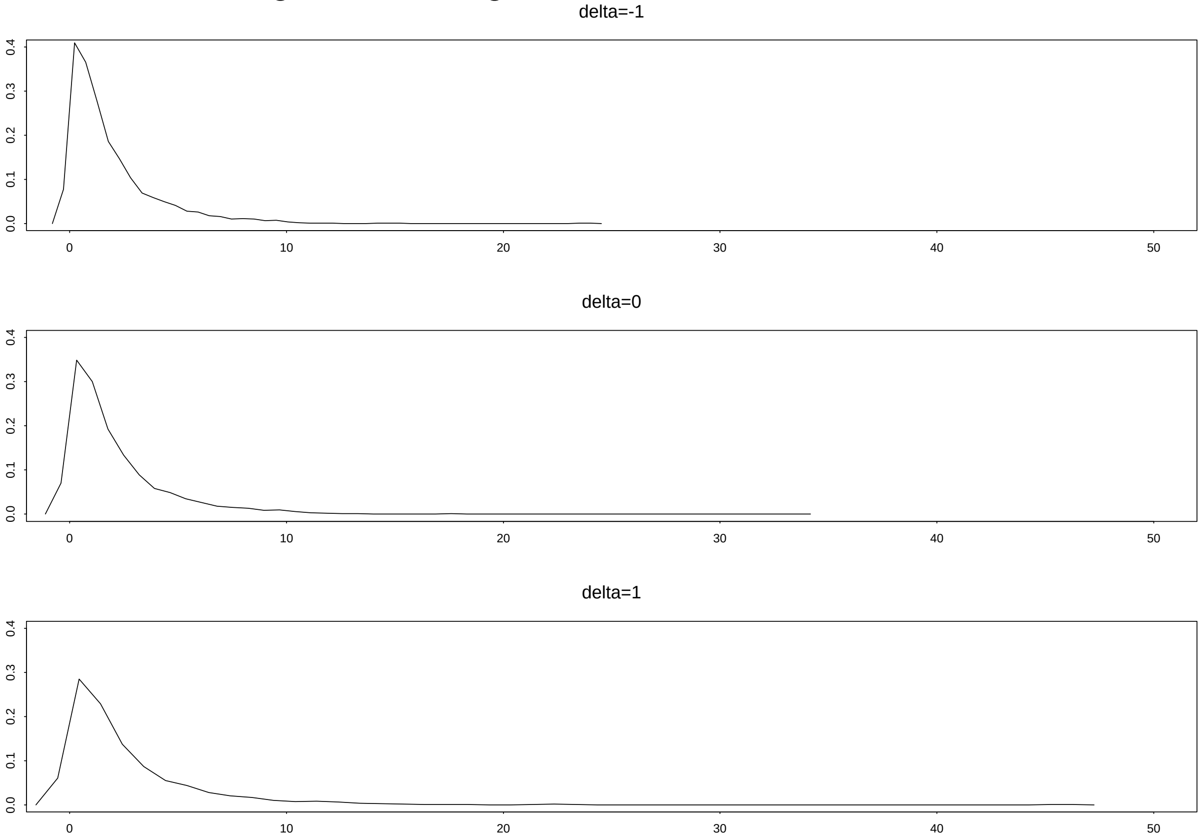


Figure 5.4: Mean Deviation from the Benchmark

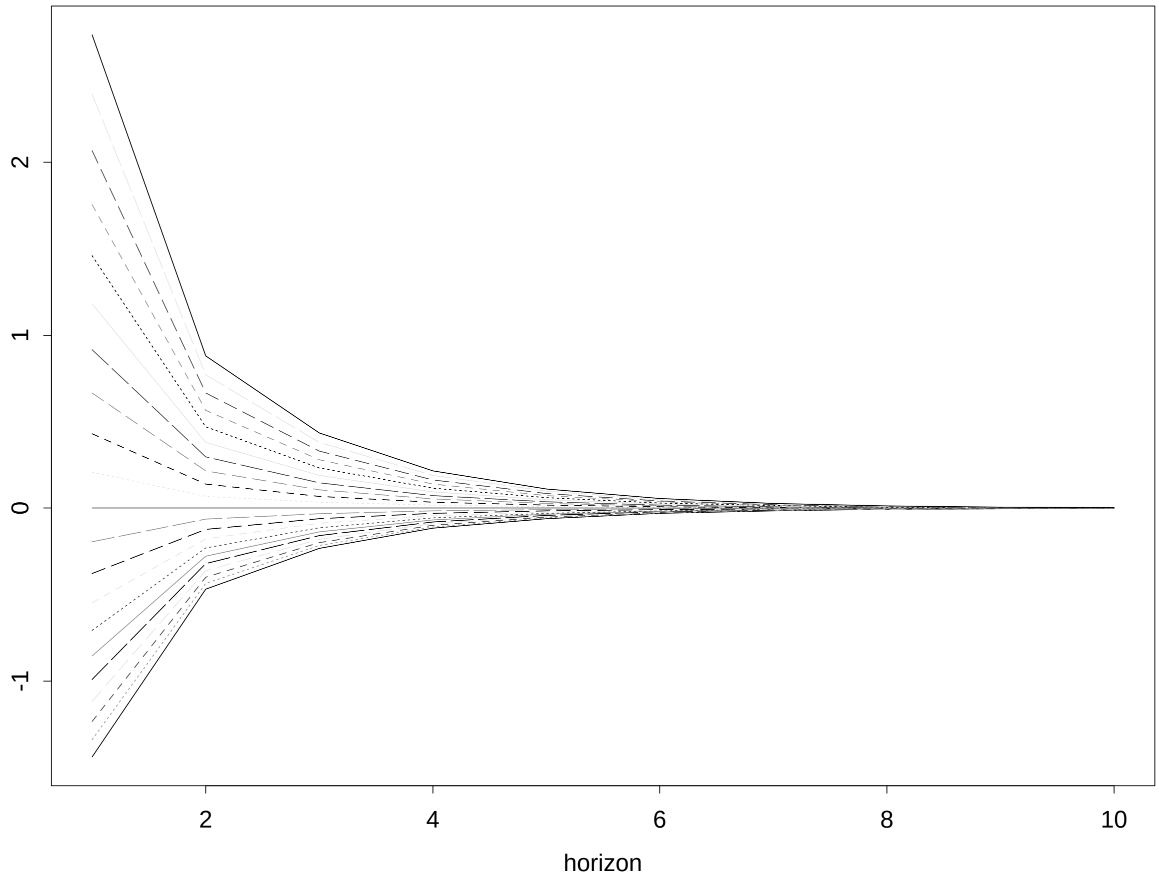


Figure 5.5: Variance of the Deviation from Benchmark

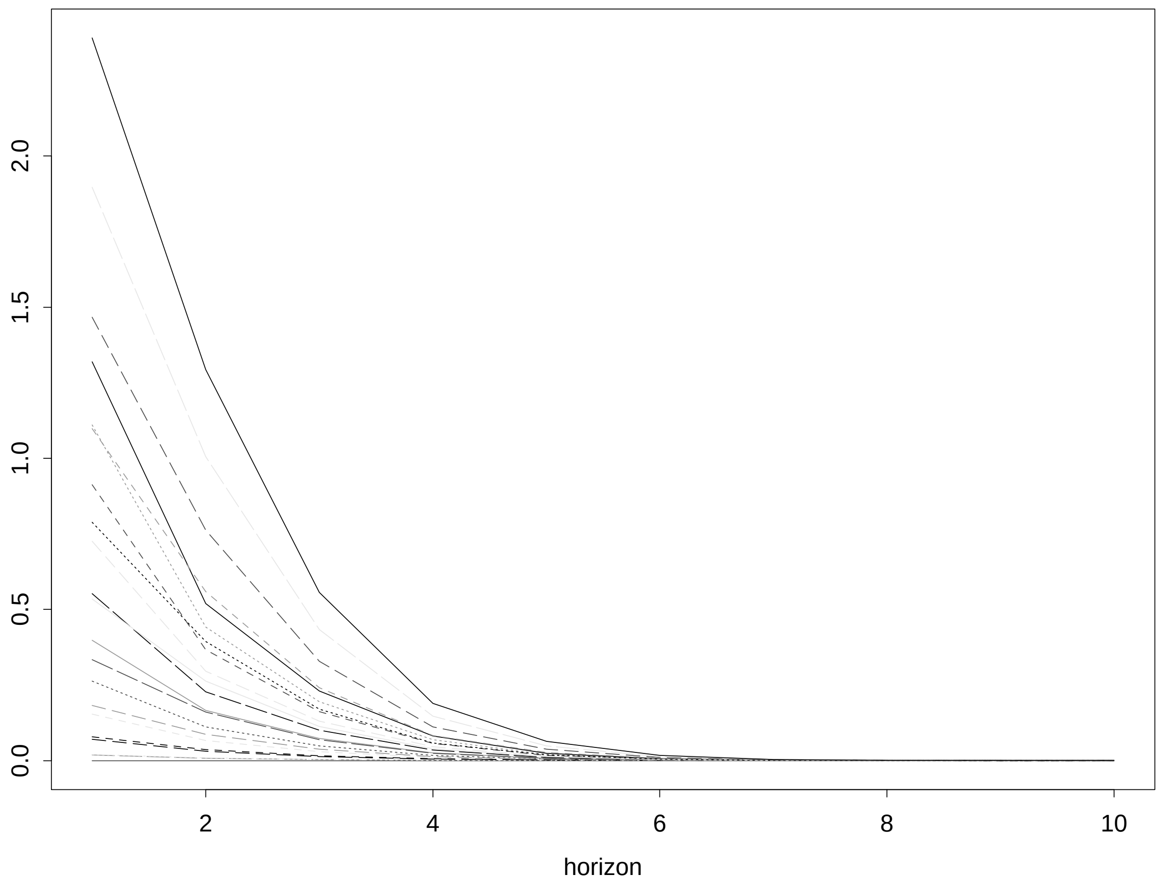


Figure 5.6: Mean Deviation from the Benchmark

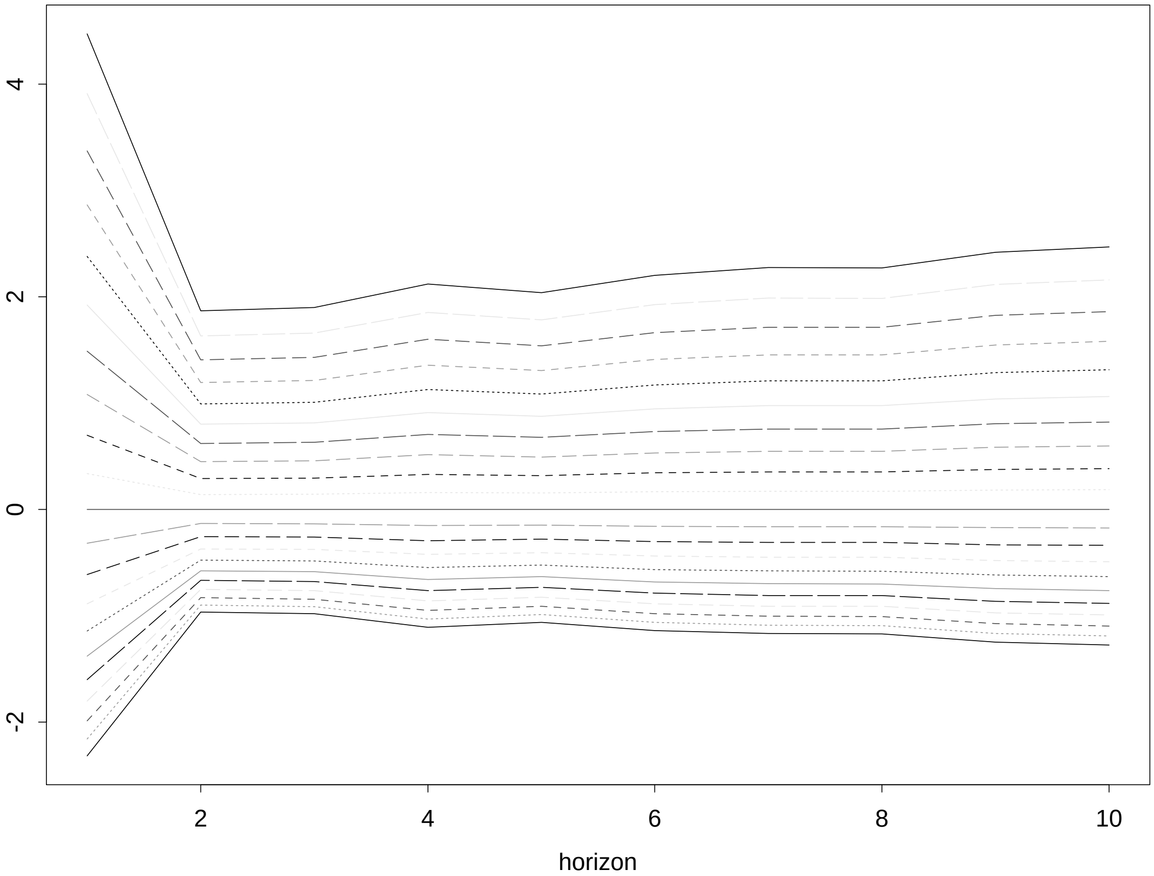


Figure 5.7: Variance of the Deviation from Benchmark

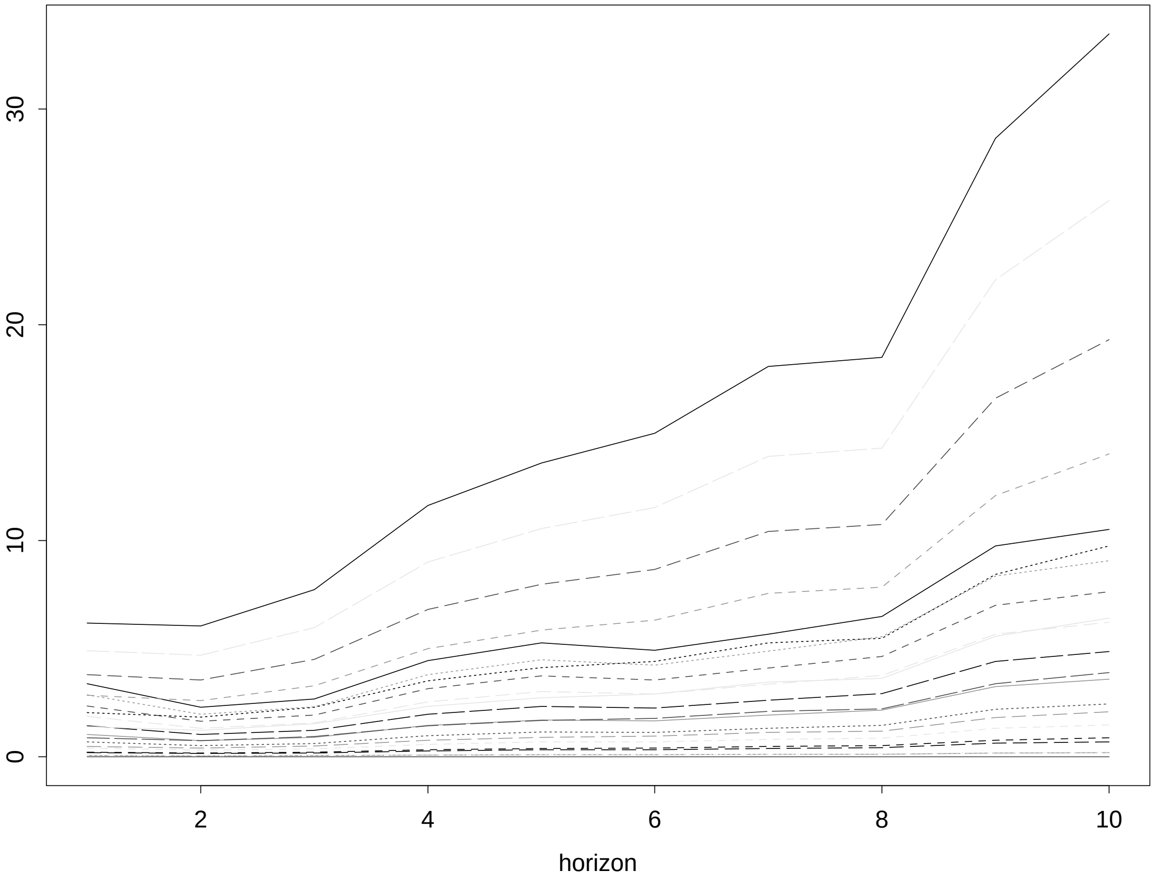


Figure 5.8: Mean Deviation from the Benchmark

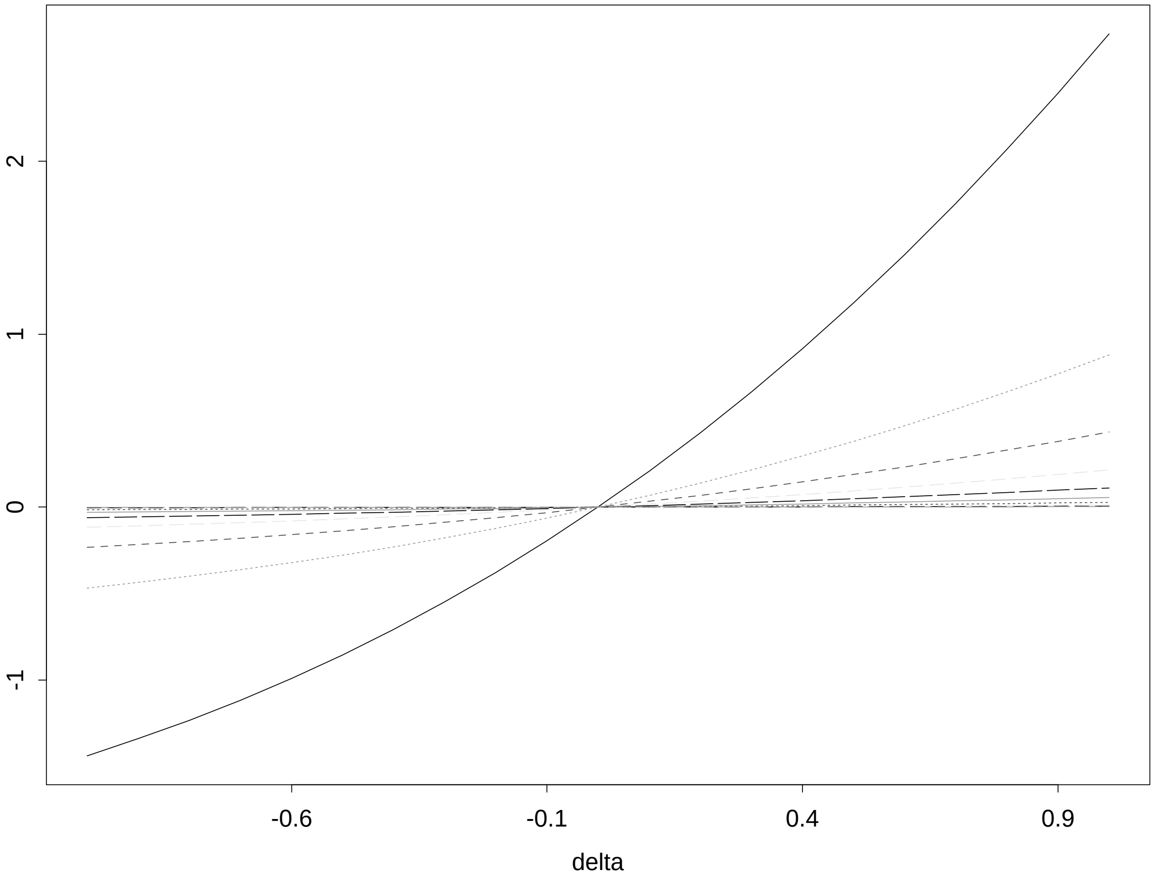


Figure 5.9: Variance of the Deviation from Benchmark

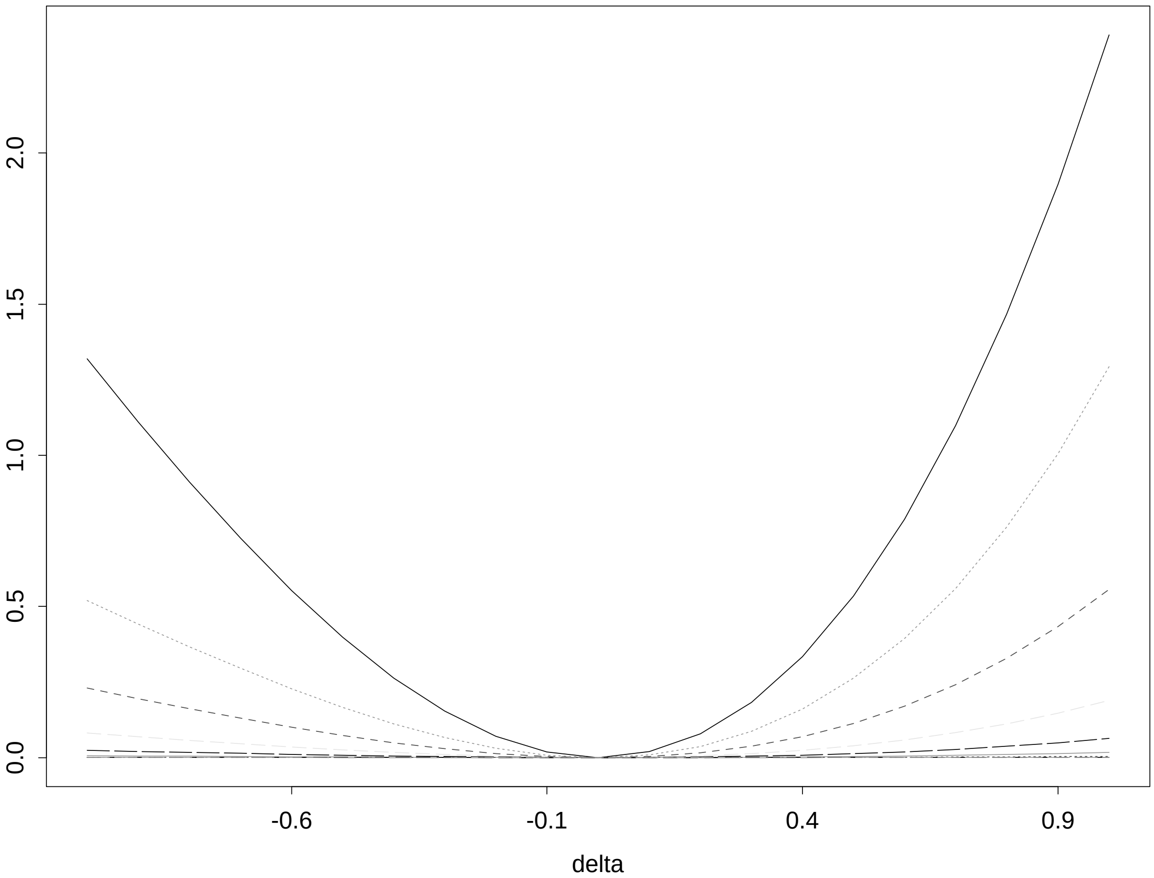


Figure 5.10: Scatterplot at Horizon 3

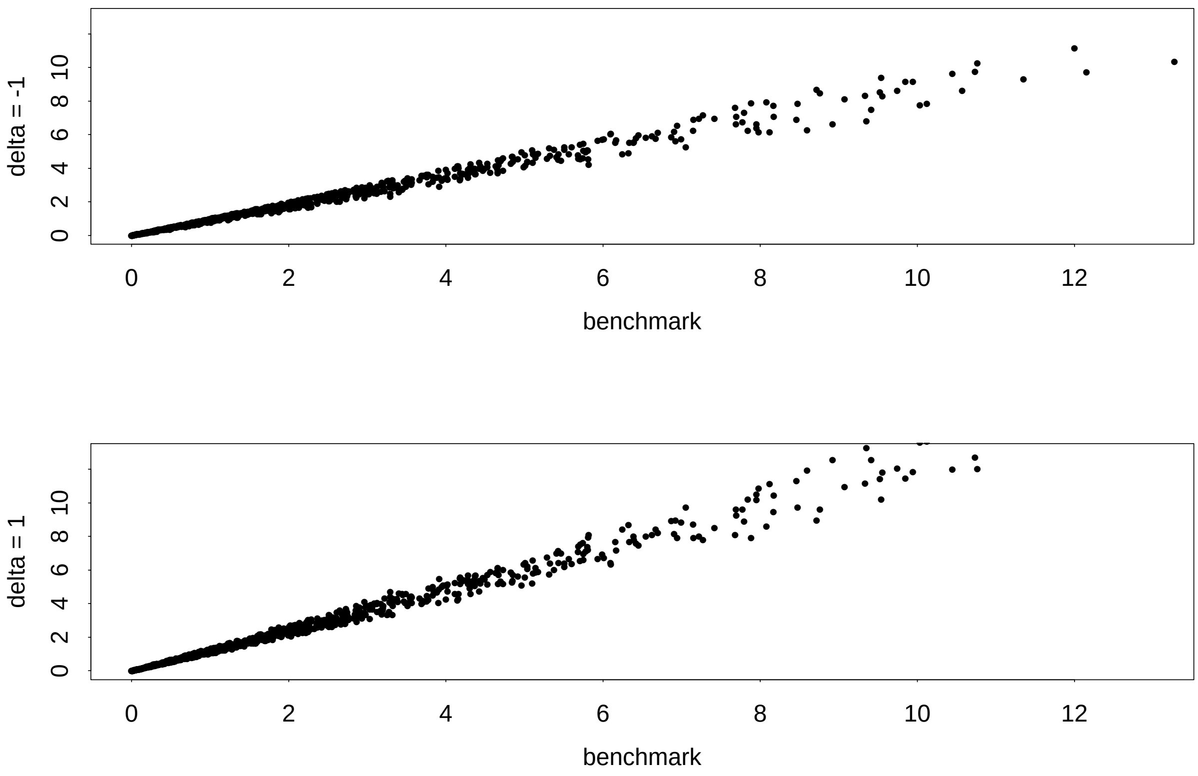


Figure 5.11: Scatterplot a = epsilon*z

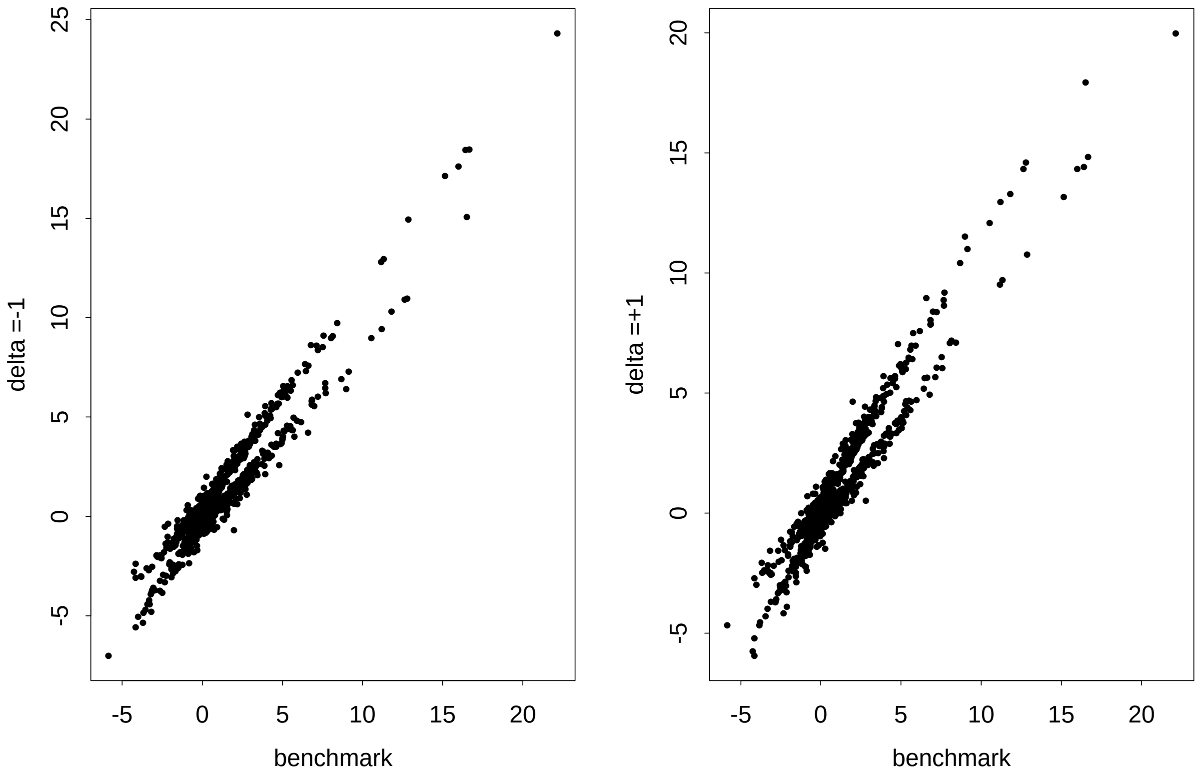
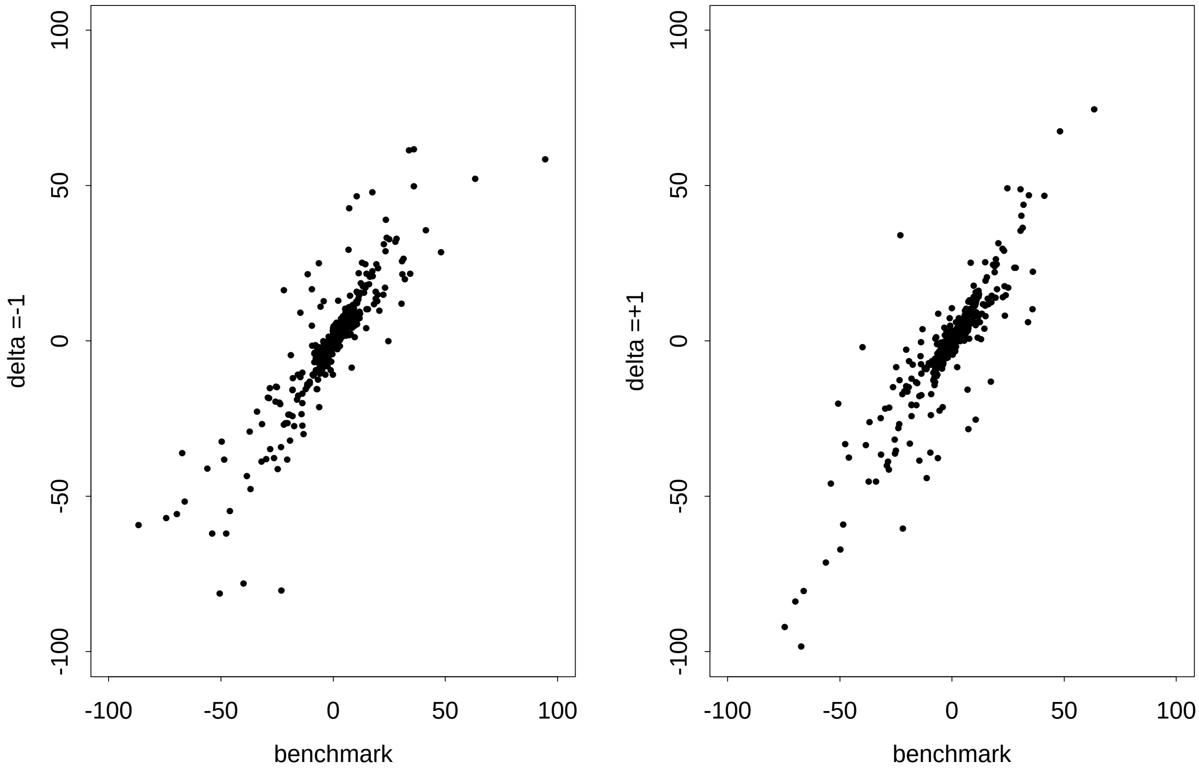


Figure 5.12: Scatterplot a = z / epsilon



References

[1] Abramowitz, M. and T. Stegun (1964) : "Handbook of Mathematical Functions", National Bureau of Standards, Applied Mathematical Series.

[2] Bay, J. (1997) : "Testing Parametric Conditional Distributions of Dynamic Models", Department of Economics, MIT.

[3] Blanchard, O. and D. Quah (1989) : "The Dynamic Effects of Aggregate Demand and Aggregate Supply Disturbances", American Economic Review, 79, 655-673.

[4] Bosq, D. (1998) : "Nonparametric Statistics for Stochastic Processes" Notes in Statistics, Springer-Verlag, New-York.

[5] Bosq, D., and D. Guegan (1995) : "Nonparametric Estimation of the Chaotic Function and the Invariant Measure of a Dynamical System", Statistics and Probability Letters, 25, 201-212.

[6] Bougerol, P., and N. Picard (1992): "Stationarity of GARCH Processes", Journal of Econometrics, 52, 115-127.

[7] Buja, A. (1990): "Remarks on Functional Canonical Variates Alternating Least Squares Methods and ACE", Annals of Statistics, 18, 1032-1069.

[8] Diebold, F, Gunther, T, and A. Tay (1997): "Evaluating Density Forecasts", manuscript, Department of Economics , University of Pennsylvania.

[9] Engle, R. (1982): "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation", Econometrica, 50, 987-1008.

[10] Engle, R., and J.R. Russell (1998): "The Autoregressive Conditional Duration Model", Econometrica, 66, 1127-1163.

[11] Gallant, A., Rossi, P. , and G. Tauchen (1993): " Nonlinear Dynamic Structures", Econometrica, 61, 871-908.

[12] Gourieroux, C., and A. Monfort (1997) : "Time Series and Dynamic Models", Cambridge University Press.

[13] Granger, C. (1995) : "Modelling Nonlinear Relationships between Extended Memory Variables", Econometrica, 63, 265-279.

[14] Granger, C., and P. Newbold (1976): "Forecasting Transformed Series", Journal of the Royal Statistical Society, B, 38, 189-203.

[15] Granger, C., and T. Terasvirta (1993) : "Modelling Nonlinear Economic Relationships", Oxford University Press.

[16] Inoue, A. (1997): " A Conditional Goodness of Fit Test in Time Series", manuscript, Department of Economics, University of Pennsylvania.

[17] Koop, G., Pesaran, H., and S. Potter (1996) : "Impulse Response Analysis in Nonlinear Multivariate Models", Journal of Econometrics, 74, 119-147.

[18] Lancaster, H. (1958) : "The Structure of Bivariate Distributions", Annals of Mathematical Statistics, 29, 719-736.

[19] Lawrance, A. and P. Lewis (1985) : "Modelling and Residual Analysis of Nonlinear Autoregressive Time Series in Exponential Variables ", J.R.S.S., B47, 165-202.

[20] Lutkepohl, H. and P. Saikkonen (1997) : "Impulse Response Analysis in Infinite Order Cointegrated Vector Autoregressive Processes", Journal of Econometrics, 81, 127-157.

[21] McLeod, A. and W. Li (1983): " Diagnostic Checking ARMA Time Series Models Using Squared Residual Autocorrelations", Journal of Time Series Analysis, 4, 269-273.

[22] Morgan, J.P. (1994) : "Risk Metrics: Technical Document", 2nd edition, New York.

[23] Nelson, D. (1990): "Stationarity and Persistence in the GARCH(1,1) Model", Econometric Theory, 6, 318-334.

[24] Nisio, M. (1960): "On Polynomial Approximation for Strictly Stationary Processes", Journal of Math. Soc. Jpn, 12, 207-276.

[25] Oseledec, V. (1968) : "A Multiplicative Ergodic Theorem: Liapunov Characteristic Numbers for Dynamical Systems" , Trans. Moscow Math. Soc., 19, 197-231.

[26] Priestley, H. (1988): " Non-linear and Non-stationary Time Series Analysis", Academic Press.

[27] Sims, C. (1972): "Money, Income and Causality", American Economic Review, 62, 540-552.

[28] Sims, C. (1980): "Macroeconomics and Reality", Econometrica , 48, 1-48.

[29] Tong, H. (1990): "Nonlinear Time Series: A Dynamic System Approach", Oxford University Press.

[30] Volterra, V. (1930): "Theory of Functionals", Blackie, London

[31] Volterra, V. (1959): " Theory of Functionals and of Integro-differential Equations" Dover, New-York.

[32] Whittle, P. (1963): "Prediction and Regulation", English University Press, London.

[33] Wiener, N. (1958): "Non-linear Problems in Random Theory", MIT Press, Cambridge, Massachussetts.

[34] Yao, Q. and H. Tong (1994) : "Quantifying the Influence of Initial Values on Nonlinear Prediction", Journal of the Royal Statistical Society , B 56, 701-725.