I. Équations différentielles stochastiques rétrogrades

L'objectif de ce chapitre est d'introduire la notion d'équations différentielles stochastiques rétrogrades, EDSR en abrégé, et de préciser la terminologie employée dans ce contexte. Nous montrerons le résultat classique d'existence et d'unicité ; des exemples d'EDSR sont donnés à la fin du chapitre.

1. Vocabulaire et notations

1.1. Présentation du problème

Considérons, sur un espace probabilisé filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, une variable aléatoire ξ mesurable par rapport à $\mathcal{F}_T$. On voudrait résoudre l'équation différentielle suivante :

$$\frac{-dY_t}{dt} = f(Y_t), \quad t \in [0,T], \qquad \text{avec}, \quad Y_T = \xi,$$

en imposant que, pour tout instant t, Y_t ne dépende pas du futur après t c'est à dire que le processus Y soit adapté par rapport à la filtration $\{\mathcal{F}_t\}_{t\geq 0}$.

Prenons l'exemple le plus simple à savoir $f \equiv 0$. Le candidat naturel est $Y_t = \xi$ qui n'est pas adapté si ξ n'est pas déterministe. La meilleure approximation – disons dans L^2 – adaptée est la martingale $Y_t = \mathbb{E}(\xi \,|\, \mathcal{F}_t)$. Si on travaille avec la filtration naturelle d'un mouvement brownien, le théorème de représentation des martingales browniennes – Théorème I.2 – permet de construire un processus Z de carré intégrable et adapté tel que :

$$Y_t = \mathbb{E}(\xi \mid \mathcal{F}_t) = \mathbb{E}[\xi] + \int_0^t Z_s \, dW_s.$$

Un calcul élémentaire montre alors que

$$Y_t = \xi - \int_t^T Z_s \, dW_s, \qquad i.e. \quad -dY_t = -Z_t \, dW_t, \quad \text{avec}, \quad Y_T = \xi.$$

On voit donc apparaître sur l'exemple le plus simple une seconde inconnue qui est le processus Z dont le rôle est de rendre le processus Y adapté.

Par conséquent, comme une seconde variable apparaît, pour obtenir la plus grande généralité, on permet à f de dépendre du processus Z ; l'équation devient donc :

$$-dY_t = f(t, Y_t, Z_t) \, dt - Z_t \, dW_t, \qquad \text{avec}, \quad Y_T = \xi.$$

1.2. Notations

On se donne $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité complet et W un MB d–dimensionnel sur cet espace. On notera $\{\mathcal{F}_t\}_{t\geq 0}$ la filtration naturelle du MB W. On travaillera avec deux espaces de processus :

- on notera tout d'abord $\mathcal{S}^2(\mathbb{R}^k)$ l'espace vectoriel formé des processus Y, progressivement mesurables, à valeurs dans $\mathbb{R}^k$, tels que :

$$\|Y\|^2_{\mathcal{S}^2} := \mathbb{E}\left[\sup_{0\leq t\leq T} |Y_t|^2\right] < \infty,$$

et $\mathcal{S}^2_c(\mathbb{R}^k)$ le sous-espace formé par les processus continus. Deux processus indistinguables seront toujours identifiés et nous garderons les mêmes notations pour les espaces quotients.

– et ensuite $\mathcal{M}^2(\mathbb{R}^{k\times d})$ celui formé par les processus Z, progressivement mesurables, à valeurs dans $\mathbb{R}^{k\times d}$, tels que :

$$\|Z\|^2_{\mathcal{M}^2} := \mathbb{E}\left[\int_0^T \|Z_t\|^2 dt\right] < \infty,$$

où si $z \in \mathbb{R}^{k\times d}$, $\|z\|^2 = \mathrm{trace}(zz^*)$. $\mathrm{M}^2(\mathbb{R}^{k\times d})$ désigne l'ensemble des classes d'équivalence de $\mathcal{M}^2(\mathbb{R}^{k\times d})$.

$\mathbb{R}^k$ et $\mathbb{R}^{k\times d}$ seront souvent omis ; les espaces $\mathcal{S}^2$, $\mathcal{S}^2_c$ et M^2 sont des espaces de Banach pour les normes définies précédemment. Nous désignerons $\mathcal{B}^2$ l'espace de Banach $\mathcal{S}^2_c(\mathbb{R}^k) \times \mathrm{M}^2(\mathbb{R}^{k\times d})$.

Dans tout ce chapitre, ainsi que dans le suivant, nous nous donnons une application aléatoire f définie sur $[0,T]\times\Omega\times\mathbb{R}^k\times\mathbb{R}^{k\times d}$ à valeurs dans $\mathbb{R}^k$ telle que, pour tout $(y,z)\in\mathbb{R}^k\times\mathbb{R}^{k\times d}$, le processus $\{f(t,y,z)\}_{0\le t\le T}$ soit progressivement mesurable. On considère également une variable aléatoire ξ, mesurable par rapport à $\mathcal{F}_T$ et à valeurs dans $\mathbb{R}^k$.

Dans ce contexte, on veut résoudre l'équation différentielle stochastique rétrograde (EDSR en abrégé) suivante :

$$-dY_t = f(t,Y_t,Z_t)\,dt - Z_t\,dW_t, \quad 0\le t\le T, \qquad Y_T = \xi,$$

ou, de façon équivalente, sous forme intégrale,

$$Y_t = \xi + \int_t^T f(r,Y_r,Z_r)\,dr - \int_t^T Z_r\,dW_r, \qquad 0\le t\le T. \tag{1}$$

La fonction f s'appelle le générateur de l'EDSR et ξ la condition terminale. Sans plus tarder, précisons ce que l'on entend par solution de l'EDSR (1).

Définition 1. Une solution de l'EDSR (1) est un couple de processus $\{(Y_t,Z_t)\}_{0\le t\le T}$ vérifiant :

1. Y et Z sont progressivement mesurables à valeurs respectivement dans $\mathbb{R}^k$ et $\mathbb{R}^{k\times d}$;
2. $\mathbb{P}$–p.s. $\displaystyle\int_0^T \left\{|f(r,Y_r,Z_r)| + \|Z_r\|^2\right\} dr < \infty$;
3. $\mathbb{P}$–p.s., on a :

$$Y_t = \xi + \int_t^T f(r,Y_r,Z_r)\,dr - \int_t^T Z_r\,dW_r, \qquad 0\le t\le T.$$

Remarque. Il est important de retenir les deux points suivants : tout d'abord, les intégrales de l'équation (1) étant bien définies, Y est une semi-martingale continue ; ensuite, comme le processus Y est progressivement mesurable, il est adapté et donc en particulier Y_0 est une quantité déterministe.

Avant de donner un premier théorème d'existence et d'unicité, nous allons montrer, que sous une hypothèse relativement faible sur le générateur f, le processus Y appartient à $\mathcal{S}^2$.

Proposition 2. *Supposons qu'il existe un processus $\{f_t\}_{0\le t\le T}$, positif, appartenant à $\mathrm{M}^2(\mathbb{R})$ et une constante positive λ tels que*

$$\forall (t,y,z)\in[0,T]\times\mathbb{R}^k\times\mathbb{R}^{k\times d}, \qquad |f(t,y,z)| \le f_t + \lambda\left(|y| + \|z\|\right).$$

Si $\{(Y_t,Z_t)\}_{0\le t\le T}$ est une solution de l'EDSR (1) telle que $Z\in\mathrm{M}^2$ alors Y appartient à $\mathcal{S}^2_c$.

Démonstration. Le résultat se déduit principalement du lemme de Gronwall et du fait que Y_0 est déterministe. En effet, on a, pour tout $t \in [0,T]$,

$$Y_t = Y_0 - \int_0^t f(r, Y_r, Z_r)\, dr + \int_0^t Z_r\, dW_r,$$

et par suite, utilisant l'hypothèse sur f,

$$|Y_t| \leq |Y_0| + \int_0^T (f_r + \lambda \|Z_r\|)\, dr + \sup_{0\leq t\leq T} \left| \int_0^t Z_r\, dW_r \right| + \lambda \int_0^t |Y_r|\, dr.$$

Posons

$$\zeta = |Y_0| + \int_0^T (f_r + \lambda \|Z_r\|)\, dr + \sup_{0\leq t\leq T} \left| \int_0^t Z_r\, dW_r \right|.$$

Par hypothèse, Z appartient à M^2 et donc, via l'inégalité de Doob, le troisième terme est de carré intégrable ; il en est de même pour $\{f_t\}_{0\leq t\leq T}$, et Y_0 est déterministe donc de carré intégrable ; il s'en suit que ζ est une variable aléatoire de carré intégrable.

Comme Y est un processus continu – cf. remarque précédente, le lemme de Gronwall fournit l'inégalité $\sup_{0\leq t\leq T} |Y_t| \leq \zeta e^{\lambda T}$ qui montre que Y appartient à $\mathcal{S}^2$. □

Remarque. Le résultat est encore valable lorsque $\|f_\cdot\|_1$ est une variable aléatoire de carré intégrable.

Finissons par un résultat d'intégrabilité qui nous servira à plusieurs reprise.

Lemme 3. *Soient $Y \in \mathcal{S}^2(\mathbb{R}^k)$ et $Z \in \mathrm{M}^2(\mathbb{R}^{k\times d})$. Alors $\left\{ \int_0^t Y_s \cdot Z_s\, dW_s,\ t \in [0,T] \right\}$ est une martingale uniformément intégrable.*

Démonstration. Les inégalités BDG donnent

$$\begin{aligned} \mathbb{E}\left[\sup_{0\leq t\leq T} \left| \int_0^t Y_r \cdot Z_r\, dW_r \right| \right] &\leq C\, \mathbb{E}\left[\left(\int_0^T |Y_r|^2 \, \|Z_r\|^2\, dr \right)^{1/2} \right] \\ &\leq C\, \mathbb{E}\left[\sup_{0\leq t\leq T} |Y_t| \left(\int_0^T \|Z_r\|^2\, dr \right)^{1/2} \right], \end{aligned}$$

et par suite, comme $ab \leq a^2/2 + b^2/2$,

$$\mathbb{E}\left[\sup_{0\leq t\leq T} \left| \int_0^t Y_r \cdot Z_r\, dW_r \right| \right] \leq C' \left(\mathbb{E}\left[\sup_{0\leq t\leq T} |Y_t|^2 \right] + \mathbb{E}\left[\int_0^T \|Z_r\|^2\, dr \right] \right).$$

Or cette dernière quantité est finie par hypothèse ; d'où le résultat. □

2. Le cas Lipschitz

2.1. Le résultat de Pardoux–Peng

Dans ce paragraphe, nous allons montrer un premier résultat d'existence et d'unicité qui sera généralisé au chapitre suivant. Ce résultat est dû à E. Pardoux et S. Peng [PP90] ; c'est le premier résultat d'existence et d'unicité pour les EDSR dans le cas où le générateur est non-linéaire.

Rappelons pour la dernière fois que f est définie sur $[0,T] \times \Omega \times \mathbb{R}^k \times \mathbb{R}^{k\times d}$ à valeurs dans $\mathbb{R}^k$, telle que, pour tout $(y,z) \in \mathbb{R}^k \times \mathbb{R}^{k\times d}$, le processus $\{f(t,y,z)\}_{0\le t\le T}$ soit progressivement mesurable. On considère également ξ une variable aléatoire, $\mathcal{F}_T$–mesurable, à valeurs dans $\mathbb{R}^k$.

Voici les hypothèses sous lesquelles nous allons travailler.

(L) Il existe une constante λ telle que $\mathbb{P}$–p.s.,

1. condition de Lipschitz en (y,z) : pour tout t, y, y', z, z',

$$\left|f(t,y,z) - f(t,y',z')\right| \le \lambda \left(|y-y'| + \|z-z'\|\right);$$

2. condition d'intégrabilité :

$$\mathbb{E}\left[|\xi|^2 + \int_0^T |f(r,0,0)|^2\, dr\right] < \infty.$$

Nous commençons par un cas très simple, celui où f ne dépend ni de y ni de z i.e. on se donne ξ de carré intégrable et un processus $\{F_t\}_{0\le t\le T}$ dans $\mathrm{M}^2(\mathbb{R}^k)$ et on veut trouver une solution de l'EDSR

$$Y_t = \xi + \int_t^T F_r\, dr - \int_t^T Z_r\, dW_r, \qquad 0 \le t \le T. \tag{2}$$

Lemme 1. *Soient $\xi \in \mathrm{L}^2(\mathcal{F}_T)$ et $\{F_t\}_{0\le t\le T} \in \mathrm{M}^2(\mathbb{R}^k)$. L'EDSR (2) possède une unique solution (Y,Z) telle que $Z \in \mathrm{M}^2$.*

Démonstration. Supposons dans un premier temps que (Y,Z) soit une solution vérifiant $Z \in \mathrm{M}^2$. Si on prend l'espérance conditionnelle sachant $\mathcal{F}_t$, on a nécessairement,

$$Y_t = \mathbb{E}\left(\xi + \int_t^T F_r\, dr \,\middle|\, \mathcal{F}_t\right).$$

On définit donc Y à l'aide de la formule précédente et il reste à trouver Z. Remarquons que, d'après le théorème de Fubini, comme F est progressivement mesurable, $\int_0^t F_r\, dr$ est un processus adapté à la filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$; en fait dans $\mathcal{S}^2_c$ puisque F est de carré intégrable. On a alors, pour tout $t \in [0,T]$,

$$Y_t = \mathbb{E}\left(\xi + \int_0^T F_r\, dr \,\middle|\, \mathcal{F}_t\right) - \int_0^t F_r\, dr := M_t - \int_0^t F_r\, dr.$$

M est une martingale brownienne ; via le Théorème I.2 on construit un processus Z appartenant à M^2 tel que

$$Y_t = M_t - \int_0^t F_r\, dr = M_0 + \int_0^t Z_r\, dW_r - \int_0^t F_r\, dr.$$

On vérifie facilement que (Y,Z) ainsi construit est une solution de l'EDSR étudiée puisque comme $Y_T = \xi$,

$$\begin{aligned} Y_t - \xi &= M_0 + \int_0^t Z_r\, dW_r - \int_0^t F_r\, dr - \left(M_0 + \int_0^T Z_r\, dW_r - \int_0^T F_r\, dr\right) \\ &= \int_t^T F_r\, dr - \int_t^T Z_r\, dW_r. \end{aligned}$$

L'unicité est évidente pour les solutions vérifiant $Z \in \mathrm{M}^2$. □

Nous montrons à présent le théorème de Pardoux et Peng.

Théorème 2. Pardoux–Peng 90. *Sous l'hypothèse (L), l'EDSR (1) possède une unique solution (Y, Z) telle que $Z \in \mathrm{M}^2$.*

Démonstration. Nous utilisons un argument de point fixe dans l'espace de Banach $\mathcal{B}^2$ en construisant une application Ψ de $\mathcal{B}^2$ dans lui-même de sorte que $(Y, Z) \in \mathcal{B}^2$ est solution de l'EDSR (1) si et seulement si c'est un point fixe de Ψ.

Pour (U, V) élément de $\mathcal{B}^2$, on définit $(Y, Z) = \Psi(U, V)$ comme étant la solution de l'EDSR :

$$Y_t = \xi + \int_t^T f(r, U_r, V_r)\, dr - \int_t^T Z_r\, dW_r, \qquad 0 \le t \le T.$$

Remarquons que cette dernière EDSR possède une unique solution qui est dans $\mathcal{B}^2$. En effet, posons $F_r = f(r, U_r, V_r)$. Ce processus appartient à M^2 puisque, f étant Lipschitz,

$$|F_r| \le |f(r, 0, 0)| + \lambda |U_r| + \lambda \|V_r\|,$$

et ces trois derniers processus sont de carré intégrable. Par suite, nous pouvons appliquer le Lemme 1 pour obtenir une unique solution (Y, Z) telle que $Z \in \mathrm{M}^2$. (Y, Z) appartient à $\mathcal{B}^2$: l'intégralité de Z est obtenue par construction et, d'après la Proposition 2, Y appartient à $\mathcal{S}^2_c$. L'application Ψ de $\mathcal{B}^2$ dans lui-même est donc bien définie.

Soient (U, V) et (U', V') deux éléments de $\mathcal{B}^2$ et $(Y, Z) = \Psi(U, V)$, $(Y', Z') = \Psi(U', V')$. Notons $y = Y - Y'$ et $z = Z - Z'$. On a, $y_T = 0$ et

$$dy_t = -\left\{ f(t, U_t, V_t) - f(t, U'_t, V'_t) \right\} dt + z_t\, dW_t.$$

On applique la formule de Itô à $e^{\alpha t}|y_t|^2$ pour obtenir :

$$d\left(e^{\alpha t}|y_t|^2\right) = \alpha e^{\alpha t}|y_t|^2\, dt - 2e^{\alpha t} y_t \cdot \left\{ f(t, U_t, V_t) - f(t, U'_t, V'_t) \right\} dt + 2e^{\alpha t} y_t \cdot z_t\, dW_t + e^{\alpha t}\|z_t\|^2\, dt.$$

Par conséquent, intégrant entre t et T, on obtient

$$\begin{aligned} e^{\alpha t}|y_t|^2 + \int_t^T e^{\alpha r}\|z_r\|^2\, dr &= \int_t^T e^{\alpha r}\left(-\alpha|y_r|^2 + 2y_r \cdot \left\{ f(r, U_r, V_r) - f(r, U'_r, V'_r) \right\}\right) dr \\ &\quad - \int_t^T 2e^{\alpha r} y_r \cdot z_r\, dW_r, \end{aligned}$$

et, comme f est Lipschitz, il vient, notant u et v pour $U - U'$ et $V - V'$ respectivement,

$$e^{\alpha t}|y_t|^2 + \int_t^T e^{\alpha r}\|z_r\|^2\, dr \le \int_t^T e^{\alpha r}\left(-\alpha|y_r|^2 + 2\lambda|y_r|\,|u_r| + 2\lambda|y_r|\,\|v_r\|\right) dr - \int_t^T 2e^{\alpha r} y_r \cdot z_r\, dW_r.$$

Pour tout $\varepsilon > 0$, on a $2ab \le a^2/\varepsilon + \varepsilon b^2$, et donc, l'inégalité précédente donne

$$\begin{aligned} e^{\alpha t}|y_t|^2 + \int_t^T e^{\alpha r}\|z_r\|^2\, dr &\le \int_t^T e^{\alpha r}\left(-\alpha + 2\lambda^2/\varepsilon\right)|y_r|^2\, dr - \int_t^T 2e^{\alpha r} y_r \cdot z_r\, dW_r \\ &\quad + \varepsilon \int_t^T e^{\alpha r}\left(|u_r|^2 + \|v_r\|^2\right) dr, \end{aligned}$$

et prenant $\alpha = 2\lambda^2/\varepsilon$, on a, notant $R_\varepsilon = \varepsilon \int_0^T e^{\alpha r}\left(|u_r|^2 + \|v_r\|^2\right) dr$,

$$\forall t \in [0, T], \qquad e^{\alpha t}|y_t|^2 + \int_t^T e^{\alpha r}\|z_r\|^2\, dr \le R_\varepsilon - 2\int_t^T e^{\alpha r} y_r \cdot z_r\, dW_r. \tag{3}$$

D'après le Lemme 3, la martingale locale $\left\{\int_0^t e^{\alpha r} y_r \cdot z_r \, dW_r\right\}_{t\in[0,T]}$ est en réalité une martingale nulle en 0 puisque Y, Y' appartiennent à $\mathcal{S}^2$ et Z, Z' appartiennent à M^2.

En particulier, prenant l'espérance – ce qui fait partir l'intégrale stochastique via la remarque précédente –, on obtient facilement, pour $t = 0$,

$$\mathbb{E}\left[\int_0^T e^{\alpha r} \|z_r\|^2 \, dr\right] \leq \mathbb{E}\left[R_\varepsilon\right]. \tag{4}$$

Revenant à l'inégalité (3), les inégalités BDG fournissent – avec C universelle –,

$$\begin{aligned}
\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|y_t|^2\right] &\leq \mathbb{E}\left[R_\varepsilon\right] + C\,\mathbb{E}\left[\left(\int_0^T e^{2\alpha r}|y_r|^2\,\|z_r\|^2\,dr\right)^{1/2}\right] \\
&\leq \mathbb{E}\left[R_\varepsilon\right] + C\,\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t/2}|y_t| \left(\int_0^T e^{\alpha r}\|z_r\|^2\,dr\right)^{1/2}\right],
\end{aligned}$$

puis, comme $ab \leq a^2/2 + b^2/2$,

$$\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|y_t|^2\right] \leq \mathbb{E}\left[R_\varepsilon\right] + \frac{1}{2}\,\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|y_t|^2\right] + \frac{C^2}{2}\,\mathbb{E}\left[\int_0^T e^{\alpha r}\|z_r\|^2\,dr\right].$$

Prenant en considération l'inégalité (4), on obtient finalement

$$\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|y_t|^2 + \int_0^T e^{\alpha r}\|z_r\|^2\,dr\right] \leq \left(3 + C^2\right)\mathbb{E}\left[R_\varepsilon\right],$$

et par suite, revenant à la définition de R_ε,

$$\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|y_t|^2 + \int_0^T e^{\alpha r}\|z_r\|^2\,dr\right] \leq \varepsilon\left(3 + C^2\right)(1\vee T)\,\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|u_t|^2 + \int_0^T e^{\alpha r}\|v_r\|^2\,dr\right].$$

Prenons ε tel que $\varepsilon(3+C^2)(1\vee T) = 1/2$, de sorte que l'application Ψ est alors une contraction stricte de $\mathcal{B}^2$ dans lui-même si on le munit de la norme

$$\|(U,V)\|_\alpha = \mathbb{E}\left[\sup_{0\leq t\leq T} e^{\alpha t}|U_t|^2 + \int_0^T e^{\alpha r}\|V_r\|^2\,dr\right]^{1/2},$$

qui en fait un espace de Banach – cette dernière norme étant équivalente à la norme usuelle correspondant au cas $\alpha = 0$.

Ψ possède donc un unique point fixe, ce qui assure l'existence et l'unicité d'une solution de l'EDSR (1) dans $\mathcal{B}^2$.

On obtient ensuite une unique solution vérifiant $Z \in \mathrm{M}^2$ puisque la Proposition 2 implique qu'un telle solution appartient à $\mathcal{B}^2$. $\square$

Remarque. À partir de maintenant et sans plus insister, l'expression « la solution de l'EDSR » signifiera la solution de l'EDSR vérifiant $Z \in \mathrm{M}^2$.

2.2. Le rôle de Z

Nous allons voir que le rôle de Z, plus précisément celui du terme $\int_t^T Z_r\,dW_r$ est de rendre le processus Y adapté et que lorsque ceci n'est pas nécessaire Z est nul.

Proposition 3. *Soit (Y,Z) la solution de l'EDSR (1) et soit τ un temps d'arrêt majoré par T. On suppose, outre l'hypothèse (L), que ξ est $\mathcal{F}_\tau$–mesurable et que $f(t,y,z)=0$ dès que $t\geq\tau$.*

Alors $Y_t=Y_{t\wedge\tau}$ et $Z_t=0$ si $t\geq\tau$.

Démonstration. On a, $\mathbb{P}$–p.s.,

$$Y_t=\xi+\int_t^T f(r,Y_r,Z_r)\,dr-\int_t^T Z_r\,dW_r,\quad 0\leq t\leq T,$$

et donc, pour $t=\tau$, comme $f(t,y,z)=0$ dès que $t\geq\tau$,

$$Y_\tau=\xi+\int_\tau^T f(r,Y_r,Z_r)\,dr-\int_\tau^T Z_r\,dW_r=\xi-\int_\tau^T Z_r\,dW_r.$$

Il vient alors $Y_\tau=\mathbb{E}(\xi|\mathcal{F}_\tau)=\xi$ et par suite $\int_\tau^T Z_r\,dW_r=0$ d'où l'on tire que

$$\mathbb{E}\left[\Big(\int_\tau^T Z_r\,dW_r\Big)^2\right]=\mathbb{E}\left[\int_\tau^T \|Z_r\|^2\,dr\right]=0,$$

et finalement que $Z_r\mathbf{1}_{r\geq\tau}=0$.

Il s'en suit immédiatement que, si $t\geq\tau$, $Y_t=Y_\tau$, puisque par hypothèse,

$$Y_\tau=Y_t+\int_\tau^t f(r,Y_r,Z_r)\,dr-\int_\tau^t Z_r\,dW_r=Y_t+0-0,$$

ce qui termine la preuve. □

Notons que dans le cas où ξ et f sont déterministes alors Z est nul et Y est la solution de l'équation différentielle

$$\frac{dY_t}{dt}=f(t,Y_t,0),\qquad Y_T=\xi.$$

2.3. Une estimation à priori

Nous finissons ce paragraphe en donnant une première estimation sur les EDSR : il s'agit en fait d'étudier la dépendance de la solution de l'EDSR par rapport aux données qui sont ξ et le processus $\{f(t,0,0)\}_{0\leq t\leq T}$. Cette étude sera reprise au chapitre suivant.

Proposition 4. *Supposons que (ξ,f) vérifie (L). Soit (Y,Z) la solution de l'EDSR (1) telle que $Z\in\mathrm{M}^2$. Alors, il existe une constante C_u universelle telle que, pour tout $\beta\geq 1+2\lambda+2\lambda^2$,*

$$\mathbb{E}\left[\sup_{0\leq t\leq T}e^{\beta t}|Y_t|^2+\int_0^T e^{\beta t}\|Z_t\|^2\,dt\right]\leq C_u\,\mathbb{E}\left[e^{\beta T}|\xi|^2+\int_0^T e^{\beta t}|f(t,0,0)|^2\,dt\right].$$

Démonstration. On applique la formule de Itô à $e^{\beta t}|Y_t|^2$ pour obtenir :

$$e^{\beta t}|Y_t|^2+\int_t^T e^{\beta r}\|Z_r\|^2\,dr=e^{\beta T}|\xi|^2+\int_t^T e^{\beta r}\left(-\beta|Y_r|^2+2Y_r\cdot f(r,Y_r,Z_r)\right)dr-\int_t^T 2e^{\beta r}Y_r\cdot Z_r\,dW_r.$$

Comme f est λ–Lipschitz, on a, pour tout (t,y,z),

$$2y\cdot f(t,y,z)\leq 2|y|\,|f(t,0,0)|+2\lambda|y|^2+2\lambda|y|\,\|z\|,$$

et donc utilisant le fait que $2ab\leq \varepsilon a^2+b^2/\varepsilon$ pour $\varepsilon=1$ puis 2,

$$2y\cdot f(t,y,z)\leq (1+2\lambda+2\lambda^2)|y|^2+|f(t,0,0)|^2+\|z\|^2/2.$$

Pour $\beta\geq 1+2\lambda+2\lambda^2$ on obtient, pour tout $t\in[0,T]$,

$$e^{\beta t}|Y_t|^2+\frac{1}{2}\int_t^T e^{\beta r}\|Z_r\|^2\,dr\leq e^{\beta T}|\xi|^2+\int_0^T e^{\beta r}|f(r,0,0)|^2\,dr-2\int_t^T e^{\beta r}Y_r\cdot Z_r\,dW_r. \tag{5}$$

La martingale locale $\left\{\int_0^t e^{\beta r}Y_r\cdot Z_r\,dW_r,\ t\in[0,T]\right\}$ est une martingale – cf. Lemme 3. En particulier, prenant l'espérance, on obtient facilement, pour $t=0$,

$$\mathbb{E}\left[\int_0^T e^{\beta r}\|Z_r\|^2\,dr\right]\leq 2\,\mathbb{E}\left[e^{\beta T}|\xi|^2+\int_0^T e^{\beta r}|f(r,0,0)|^2\,dr\right].$$

Revenant à l'inégalité (5), les inégalités BDG fournissent – avec C universelle –,

$$\mathbb{E}\left[\sup_{0\leq t\leq T}e^{\beta t}|Y_t|^2\right]\leq \mathbb{E}\left[e^{\beta T}|\xi|^2+\int_0^T e^{\beta r}|f(r,0,0)|^2\,dr\right]+C\,\mathbb{E}\left[\left(\int_0^T e^{2\beta r}|Y_r|^2\,\|Z_r\|^2\,dr\right)^{1/2}\right].$$

D'autre part,

$$\begin{aligned}C\,\mathbb{E}\left[\left(\int_0^T e^{2\beta r}|Y_r|^2\,\|Z_r\|^2\,dr\right)^{1/2}\right] &\leq C\,\mathbb{E}\left[\sup_{0\leq t\leq T}e^{\beta t/2}|Y_t|\left(\int_0^T e^{\beta r}\|Z_r\|^2\,dr\right)^{1/2}\right]\\ &\leq \frac{1}{2}\,\mathbb{E}\left[\sup_{0\leq t\leq T}e^{\beta t}|Y_t|^2\right]+\frac{C^2}{2}\,\mathbb{E}\left[\int_0^T e^{\beta r}\|Z_r\|^2\,dr\right].\end{aligned}$$

Il vient alors,

$$\mathbb{E}\left[\sup_{0\leq t\leq T}e^{\beta t}|Y_t|^2\right]\leq 2\,\mathbb{E}\left[e^{\beta T}|\xi|^2+\int_0^T e^{\beta r}|f(r,0,0)|^2\,dr\right]+C^2\,\mathbb{E}\left[\int_0^T e^{\beta r}\|Z_r\|^2\,dr\right],$$

et finalement on obtient

$$\mathbb{E}\left[\sup_{0\leq t\leq T}e^{\beta t}|Y_t|^2+\int_0^T e^{\beta r}\|Z_r\|^2\,dr\right]\leq 2\left(2+C^2\right)\,\mathbb{E}\left[e^{\beta T}|\xi|^2+\int_0^T e^{\beta r}|f(r,0,0)|^2\,dr\right],$$

ce qui termine la preuve de la proposition prenant $C_u=2\left(2+C^2\right)$. □

3. Théorème de comparaison

3.1. EDSR linéaires

Dans ce paragraphe nous étudions le cas particulier des EDSR linéaires pour lesquelles nous allons donner une formule plus ou moins explicite.

On se place dans le cas $k=1$; Y est donc réel et Z est une matrice de taille $1\times d$ c'est à dire un vecteur ligne de dimension d.

Proposition 1. *Soit* $\{(a_t, b_t)\}_{t\in[0,T]}$ *un processus à valeurs dans* $\mathbb{R}\times\mathbb{R}^d$*, progressivement mesurable et borné. Soient* $\{c_t\}_{t\in[0,T]}$ *un élément de* $\mathrm{M}^2(\mathbb{R})$ *et* ξ *une variable aléatoire,* $\mathcal{F}_T$*–mesurable, de carré intégrable, à valeurs réelles.*

L'EDSR linéaire

$$Y_t = \xi + \int_t^T \{a_r Y_r + Z_r b_r + c_r\}\, dr - \int_t^T Z_r\, dW_r,$$

possède une unique solution qui vérifie :

$$\forall t \in [0,T], \qquad Y_t = \Gamma_t^{-1}\, \mathbb{E}\left(\xi\Gamma_T + \int_t^T c_r \Gamma_r\, dr \;\Big|\; \mathcal{F}_t\right),$$

avec, pour tout $t \in [0,T]$*,*

$$\Gamma_t = \exp\left\{\int_0^t b_r \cdot dW_r - \frac{1}{2}\int_0^t |b_r|^2\, dr + \int_0^t a_r\, dr\right\}.$$

Démonstration. Commençons par remarquer que le processus Γ vérifie :

$$d\Gamma_t = \Gamma_t\,(a_t\, dt + b_t \cdot dW_t)\,, \qquad \Gamma_0 = 1.$$

D'autre part, comme b est borné, l'inégalité de Doob montre que Γ appartient à $\mathcal{S}^2$.

De plus, les hypothèses de cette proposition assure l'existence d'une unique solution (Y,Z) à l'EDSR linéaire ; il suffit de poser $f(t,y,z) = a_t y + z b_t + c_t$ et de vérifier que (L) est satisfaite. Y appartient à $\mathcal{S}^2$ par la Proposition 2.

La formule d'intégration par parties donne

$$d\Gamma_t Y_t = \Gamma_t\, dY_t + Y_t\, d\Gamma_t + d\langle \Gamma, Y\rangle_t = -\Gamma_t c_t\, dt + \Gamma_t Z_t\, dW_t + \Gamma_t Y_t b_t \cdot dW_t,$$

ce qui montre que le processus $\Gamma_t Y_t + \int_0^t c_r \Gamma_r\, dr$ est une martingale locale qui est en fait une martingale car $c \in \mathrm{M}^2$ et Γ, Y sont dans $\mathcal{S}^2$.

Par suite,

$$\Gamma_t Y_t + \int_0^t c_r \Gamma_r\, dr = \mathbb{E}\left(\Gamma_T Y_T + \int_0^T c_r \Gamma_r\, dr \;\Big|\; \mathcal{F}_t\right),$$

ce qui donne la formule annoncée. □

Remarque. Notons que si $\xi \geq 0$ et si $c_t \geq 0$ alors la solution de l'EDSR linéaire vérifie $Y_t \geq 0$. Cette remarque va nous permettre d'obtenir le théorème de comparaison au paragraphe suivant.

Pour illustrer ce résultat prenons le cas où a et c sont nuls. On a alors

$$Y_t = \mathbb{E}\left(\xi \exp\left\{\int_t^T b_r \cdot dW_r - \frac{1}{2}\int_t^T |b_r|^2\, dr\right\} \;\Big|\; \mathcal{F}_t\right) = \mathbb{E}^*\left(\xi \mid \mathcal{F}_t\right),$$

où $\mathbb{P}^*$ est la mesure de densité par rapport à $\mathbb{P}$

$$L_T = \exp\left\{\int_0^T b_r \cdot dW_r - \frac{1}{2}\int_0^T |b_r|^2\, dr\right\}.$$

Une autre façon de voir cela, plus dans l'esprit « probabilité risque neutre », est de regarder l'EDSR sous $\mathbb{P}^*$. En effet, sous $\mathbb{P}^*$, $B_t = W_t - \int_0^t b_r\, dr$ est un MB – c'est le théorème de Girsanov cf. Théorème I.3. Or l'équation peut s'écrire

$$-dY_t = Z_t b_t\, dt - Z_t\, dW_t = -Z_t\, dB_t, \qquad Y_T = \xi.$$

Donc, sous $\mathbb{P}^*$, Y est une martingale, ce qui montre aussi la formule.

On retrouve ainsi les changements de mesures de probabilité du type « transformation de Girsanov ».

3.2. Théorème de comparaison

Ce paragraphe est consacré au « théorème de comparaison » qui permet de comparer les solutions de deux EDSR (dans $\mathbb{R}$) dès que l'on sait comparer les conditions terminales et les générateurs. Ce théorème est dû à l'origine à S. PENG [Pen92].

Théorème 2. *Supposons que $k = 1$ et que (ξ, f), (ξ', f') vérifient l'hypothèse (L). On note (Y, Z) et (Y', Z') les solutions des EDSR correspondantes. On suppose également que $\mathbb{P}$–p.s. $\xi \le \xi'$ et que $f(t, Y_t, Z_t) \le f'(t, Y_t, Z_t)$ $m \otimes \mathbb{P}$–p.p. (m mesure de Lebesgue). Alors,*

$$\mathbb{P} - p.s., \qquad \forall t \in [0, T], \qquad Y_t \le Y'_t.$$

Si de plus, $Y_0 = Y'_0$, alors $\mathbb{P}$–p.s., $Y_t = Y'_t$, $0 \le t \le T$ et $f(t, Y_t, Z_t) = f'(t, Y_t, Z_t)$ $m \otimes \mathbb{P}$–p.p. En particulier, dès que $\mathbb{P}(\xi < \xi') > 0$ ou $f(t, Y_t, Z_t) < f'(t, Y_t, Z_t)$ sur un ensemble de $m \otimes \mathbb{P}$–mesure strictement positive alors $Y_0 < Y'_0$.

Démonstration. La preuve s'effectue par linéarisation ce qui permet de se ramener aux EDSR linéaires. On cherche une équation satisfaite par $U = Y' - Y$; on a notant $V = Z' - Z$ et $\zeta = \xi' - \xi$,

$$U_t = \zeta + \int_t^T \big(f'(r, Y'_r, Z'_r) - f(r, Y_r, Z_r)\big)\, dr - \int_t^T V_r\, dW_r.$$

On découpe l'accroissement des f en trois morceaux en écrivant

$$\begin{aligned} f'(r, Y'_r, Z'_r) - f(r, Y_r, Z_r) \;&=\; f'(r, Y'_r, Z'_r) - f'(r, Y_r, Z'_r) + f'(r, Y_r, Z'_r) - f'(r, Y_r, Z_r) \\ &\quad + f'(r, Y_r, Z_r) - f(r, Y_r, Z_r) \;\; \text{(qui est positif ici)}. \end{aligned}$$

On introduit deux processus a et b : a est à valeurs réelles et b est un vecteur (colonne) de dimension d. On pose :

$$a_r = \frac{f'(r, Y'_r, Z'_r) - f'(r, Y_r, Z'_r)}{U_r}, \quad \text{si} \quad U_r \ne 0, \qquad \text{et} \quad a_r = 0 \quad \text{sinon}.$$

Pour définir b, on doit introduire une autre notation : pour $0 \le i \le d$, $Z_r^{(i)}$ est la ligne dont les $d - i$ dernières composantes sont celles de Z'_r et les i premières celles de Z_r. Pour $1 \le i \le d$, on pose

$$b_r^i = \frac{f'\left(r, Y_r, Z_r^{(i-1)}\right) - f'\left(r, Y_r, Z_r^{(i)}\right)}{V_r^i}, \quad \text{si} \quad V_r^i \ne 0, \qquad \text{et} \quad b_r^i = 0 \quad \text{sinon}.$$

Remarquons que, puisque f' est Lipschitz, ces deux processus sont progressivement mesurables et bornés. Avec ces notations, on a,

$$U_t = \zeta + \int_t^T (a_r U_r + V_r b_r + c_r)\, dr - \int_t^T V_r\, dW_r,$$

où $c_r = f'(r, Y_r, Z_r) - f(r, Y_r, Z_r)$. Par hypothèse, on a $\zeta \ge 0$ et $c_r \ge 0$. Utilisant la formule « explicite » pour les EDSR linéaires – Proposition 1 –, on a, pour $t \in [0, T]$,

$$U_t = {\Gamma_t}^{-1}\, \mathbb{E}\left(\zeta \Gamma_T + \int_t^T c_r \Gamma_r\, dr \,\Big|\, \mathcal{F}_t\right),$$

avec, pour $0 \le r \le T$,

$$\Gamma_r = \exp\Big\{ \int_0^r b_u \cdot dW_u - \frac{1}{2} \int_0^r |b_u|^2\, du + \int_0^r a_u\, du \Big\}.$$

Comme déjà mentionné lors de la remarque suivant la Proposition 1, cette formule montre que $U_t \geq 0$, dès que $\zeta \geq 0$ et $c_r \geq 0$.

Pour la seconde partie du résultat, si de plus $U_0 = 0$ on a :

$$0 = \mathbb{E}\left(\zeta\Gamma_T + \int_0^T c_r\Gamma_r\, dr\right),$$

et la variable aléatoire intégrée est positive. Par conséquent, elle est nulle $\mathbb{P}$–p.s. ce qui termine la preuve de ce théorème en remarquant que dans ce cas $\zeta = 0$ et $c_r = 0$. $\square$

Remarque. On peut supposer que $f(t, Y'_t, Z'_t) \leq f'(t, Y'_t, Z'_t)$ au lieu de $f(t, Y_t, Z_t) \leq f(t, Y_t, Z_t)$ pour obtenir le résultat précédent. Il suffit de faire une linéarisation en partant de l'écriture

$$\begin{aligned} f'(r, Y'_r, Z'_r) - f(r, Y_r, Z_r) &= f'(r, Y'_r, Z'_r) - f(r, Y'_r, Z'_r) \quad \text{(supposé positif ici)} \\ &\quad + f(r, Y'_r, Z'_r) - f(r, Y_r, Z'_r) + f(r, Y_r, Z'_r) - f(r, Y_r, Z_r). \end{aligned}$$

3.3. Modèle de Black–Scholes

Le modèle de Black–Scholes conduit à un exemple d'EDSR linéaire. Prenons le cas le plus simple. On a se place dans un marché financier sur lequel on a une action dont le prix d'une part est régi par l'EDS

$$dS_t = S_t(\mu\, dt + \sigma\, dW_t), \qquad S_0 = x,$$

où $\mu \in \mathbb{R}$, $\sigma > 0$; le paramètre σ s'appelle la volatilité. On a, pour tout $t \geq 0$,

$$S_t = x \exp\left\{\sigma W_t + (\mu - \sigma^2/2)t\right\}.$$

Parallèlement à cette action, on a un placement sans risque – disons un livret de Caisse d'épargne pour fixer les idées – dont le taux de rendement est constant, égal à r ; le prix d'une part est donné part

$$dE_t = rE_t\, dt, \quad E_0 = y, \qquad \text{i.e.} \quad E_t = ye^{rt}.$$

Une stratégie est la donnée d'un couple de processus, $(p_t, q_t)_{t\geq 0}$, adapté par rapport à la filtration du MB W ; q_t représente le nombre de parts d'actif sans risque et p_t celui d'actif risqué i.e. le nombre d'actions détenues dans le portefeuille à l'instant t. La valeur du portefeuille est donc, à l'instant t,

$$V_t = q_tE_t + p_tS_t.$$

On ne considère que des stratégies autofinancées ce qui se traduit par le fait que l'évolution de la valeur du portefeuille est décrite par

$$dV_t = q_t\, dE_t + p_t\, dS_t = rq_tE_t\, dt + p_tS_t(\mu\, dt + \sigma\, dW_t).$$

Or $q_tE_t = V_t - p_tS_t$ et donc, on a, notant $\pi_t = p_tS_t$ – la somme d'argent détenue en actions –

$$dV_t = rV_t\, dt + \pi_t\sigma(\mu - r)/\sigma\, dt + \pi_t\sigma\, dW_t,$$

soit encore notant $Z_t = \pi_t\sigma$ et $\theta = (\mu - r)/\sigma$ le « risk premium »,

$$dV_t = rV_t\, dt + \theta Z_t\, dt + Z_t\, dW_t.$$

Un problème fréquent en finance consiste à donner un prix aux options. Une option européenne d'achat, un « call », de maturité T et de prix d'exercice K est un contrat qui donne le droit mais non l'obligation à son détenteur d'acheter une part de l'action au prix d'exercice K

à la date T. Le vendeur de l'option s'engage donc à payer à son détenteur la somme $(S_T - K)^+$ qui représente le profit que permet l'exercice de l'option. Plus généralement on peut imaginer un actif contingent dont le bénéfice est une variable aléatoire positive ξ qui dépend de $(S_t)_{0\leq t\leq T}$. À quel prix v vendre l'option ? Le vendeur doit s'assurer qu'en vendant l'option à ce prix à la date $t = 0$, il disposera de la somme ξ à la date $t = T$.

Pour trouver v, l'idée fondamentale est celle de duplication : le vendeur vend l'option au prix v et investit cette somme dans le marché en suivant la stratégie $(Z_t)_{0\leq t\leq T}$ à trouver ! La valeur de son portefeuille est régie par l'EDS

$$dV_t = rV_t\,dt + \theta Z_t\,dt + Z_t\,dW_t, \qquad V_0 = v.$$

Le problème est alors de trouver v et $\{Z_t\}_{0\leq t\leq T}$ de sorte que la solution de l'EDS précédente vérifie $V_T = \xi$; on dit que dans ce cas que v est le prix équitable. En d'autres termes, peut-on trouver $\{(V_t, Z_t)\}_{0\leq t\leq T}$ adapté tels que :

$$dV_t = rV_t\,dt + \theta Z_t\,dt + Z_t\,dW_t, \qquad V_T = \xi,$$

auquel cas il suffit de vendre l'option au prix $v = V_0$. Il s'agit dans ce cas de résoudre une EDSR qui est linéaire.

Supposons maintenant que « le régulateur du marché » veuille éviter la revente instantanée de l'action. Il peut soit interdire formellement cette transaction soit pénaliser les investisseurs qui se livrent à cette pratique – par exemple en leur faisant verser une somme proportionnelle à $\beta\pi_t^- = \gamma Z_t^-$ $(\gamma > 0)$. Dans ce dernier cas, dupliquer un actif contingent revient à résoudre l'EDSR,

$$dV_t = (rV_t + \theta Z_t - \gamma Z_t^-)\,dt + Z_t\,dW_t, \qquad V_T = \xi.$$

L'EDSR n'est plus linéaire mais vérifie néanmoins l'hypothèse (L). Un autre exemple d'EDSR non-linéaire intervenant en finance est le suivant :

$$dV_t = (rV_t + \theta Z_t)\,dt + Z_t\,dW_t - (R - r)(V_t - Z_t/\sigma)^-\,dt, \qquad V_T = \xi.$$

On est amené à résoudre cette dernière équation pour dupliquer un actif contingent lorsque le taux d'emprunt est $R > r$.

Signalons que dans tous les exemples précédents, les stratégies sont admissibles i.e. $V_t \geq 0$. Cela résulte du théorème de comparaison puisque $\xi \geq 0$ et $f(t, 0, 0) \geq 0$.

Références

[Pen92] S. Peng, *Stochastic Hamilton-Jacobi-Bellman equations*, SIAM J. Control Optim. **30** (1992), no. 2, 284–304.

[PP90] E. Pardoux and S. Peng, *Adapted solution of a backward stochastic differential equation*, Systems Control Lett. **14** (1990), no. 1, 55–61.