Analyse numérique : évolution
M1 math appli 2020-21
Université Paris Dauphine - PSL
Gabriel Turinici

Cours 1: introduction

$-S(t+\Delta t)+S(t)$ = # de personnes infectées entre t et $t+\Delta t$

$\downarrow$

proportionnel $S, I, \Delta t, \beta, r$

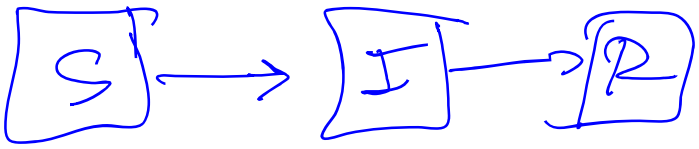
$\downarrow$

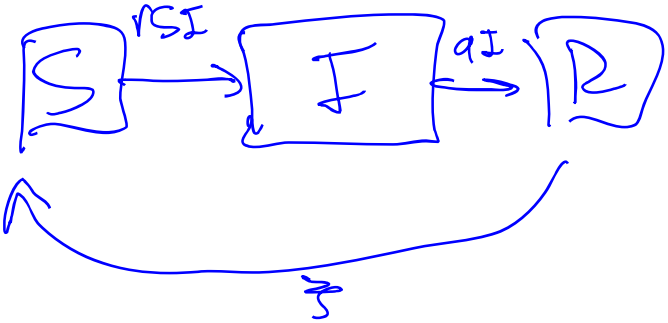
$$S(t)-S(t+\Delta t) \simeq \Delta t \cdot S(t) \cdot I(t) \cdot r$$

$$\frac{S(t)-S(t+\Delta t)}{\Delta t} \simeq S(t) \cdot I(t) \cdot r \qquad \Delta t \to 0 \quad \ldots$$

$$\boxed{S} \longrightarrow \boxed{I} \longrightarrow \boxed{R}$$

$$\begin{tikzcd} \boxed{S} \arrow[r, "rSI"] & \boxed{I} \arrow[r, "aI"] & \boxed{R} \arrow[ll, bend left, "\xi"] \end{tikzcd}$$





$$\dot{R} = aI - \xi R$$

$$\dot{S} = -rSI + \xi R$$

$$S_0 \simeq 6J \cdot 10^6$$

$$\begin{cases} \dot{S} = -\dfrac{rSI}{N} \\ \dot{I} = \dfrac{rSI}{N} - aI \\ \dot{R} = \lambda aI \qquad \dot{D} = (1-\lambda)\, aI \end{cases}$$

$$N = S+I+R$$

$$\begin{tikzcd} \boxed{S} \arrow[r] & \boxed{I} \arrow[r] \arrow[dr] & \boxed{R} \\ & & \boxed{D} \end{tikzcd}$$

$70K \sim D$
$N = 65M$

$I/N \sim$ prevalence

$R = \#$ de rétablis

$R_t =$ tx de prop r/a

EDS $\quad \underbrace{\dfrac{S_{t+\Delta t} - S_t}{S_t}}_{\sim \text{ rendement}} \sim \mathcal{N}(\mu \Delta t, \sigma^2 \Delta t)$

$\sigma^2 = 0$

$\downarrow$

$\dfrac{S_{t+\Delta t} - S_t}{S_t} = \mu \Delta t \qquad \dfrac{dS_t}{dt} = \mu S_t \qquad \dot{S} = \mu S \quad$ EDO

limite $\Delta t \to 0$

$\sigma \neq 0$ EDS $\quad S_t =$ variable aléatoire

Calcul de dérivée : question d'efficacité numérique.

Cours 2: EDO

Construction de méthodes R-K d'ordre 2 (explicites)

$s=2$ Méthode explicite $\Rightarrow A$ est triangulaire inférieure stricte $A = \begin{pmatrix} 0 & 0 \\ a_{21} & 0 \end{pmatrix} =: \begin{pmatrix} 0 & 0 \\ a & 0 \end{pmatrix}$

$a := a_{21}$. $c_1 = 0+0 = 0$ $c_2 = 0+a = a$

Tableau de Butcher

$$\begin{array}{c|cc} 0 & 0 & 0 \\ a & a & 0 \\ \hline & b_1 & b_2 \end{array}$$

Écriture équivalente :

$U_{n+1} = U_n + hF$ avec $F = b_1 K_1 + b_2 K_2$

$K_1 = f(t_n + 0, U_n + h\cdot 0) = f_n$

$K_2 = f(t_n + ah, U_n + h\cdot aK_1) = f(t_n + ah, U_n + ahf_n)$

Pour être d'ordre 2 il faut que $\varepsilon(h) = O(h^2)$.

$\tau_{n+1}(h) = \dfrac{X_{n+1} - U_{n+1}^*}{h}$ ← sol numérique partant de X_n. voir remarque 2.5.

$U_{n+1}^* = X_n + h\left[b_1 f(t_n, X_n) + b_2 f(t_n + ah, X_n + ahf(t_n, X_n))\right]$

Rappel (Taylor 2D) $G(\alpha+h_1, \beta+h_2) = G(\alpha,\beta) + \frac{\partial G}{\partial \alpha}(\alpha,\beta)\cdot h_1 + \frac{\partial G}{\partial \beta}(\alpha,\beta) h_2 + O\left[\left(\sqrt{h_1^2+h_2^2}\right)^2\right]$

Pour nous $f(t_n + ah, X_n + ahf(t_n, X_n)) = f(t_n, X_n) + \underbrace{f_t}_{\frac{\partial f}{\partial t}}(t_n, X_n)\cdot ah + f_x(t_n, X_n)\cdot ahf_n + O(h^2)$

$\frac{d}{dt}\underbrace{f(t, X(t))}_{X'(t)} = f_t(t, X(t)) + f_x(t, X(t))\cdot \underbrace{X'(t)}_{f(t, X(t))} = f_t + f_x \cdot f$

$X''(t) = f_t + f\cdot f_x$

$\partial_2 f \quad \partial_x f$

f = notation pour f_n

$X_{n+1} = X(t_n + h)$

$$\tau_{n+1}(h) = \frac{x_{n+1} - u^*_{n+1}}{h} \overset{\text{Taylor}}{=} \frac{x_{n+1} - x_n - h b_1 f - h b_2 f(t_n + ah, x_n + ah f)}{h}$$

$$= \frac{\overbrace{x'(t_n)}^{f} \cdot h + x''(t_n)\cdot\frac{h^2}{2} + O(h^3) - h b_1 f - h b_2 \left(f + f_t \cdot ah + f_x \cdot ah f\right) + O(h^3)}{h}$$

$$= f\cdot(1 - b_1 - b_2) + h \cdot \left[\frac{x''}{2} - a b_2 (f_t + f_x f)\right] + O(h^2)$$

$$= f\cdot(1 - b_1 - b_2) + h\, x''\left(\frac{1}{2} - a b_2\right) + O(h^2)$$

Pour que le schéma soit d'ordre 2 il faut (et il suffit) que

$$\begin{cases} 1 - b_1 - b_2 = 0 \\ \frac{1}{2} = a b_2 \end{cases} \iff \begin{cases} b_1 + b_2 = 1 \\ a b_2 = 1/2 \end{cases}$$

a	a	0
	$1 - \frac{1}{2a}$	$\frac{1}{2a}$

Exemple $b_1 = b_2 \Rightarrow b_1 = b_2 = 1/2 \Rightarrow a = 1$

1	1	0
	$1/2$	$1/2$

$k_1 = f_n$, $k_2 = f(t_n + h, u_n + h f_n)$ $\quad u_{n+1} = u_n + \frac{h}{2}\left(f_n + f(t_{n+1}, u_n + h f_n)\right)$

c'est le schéma de Heun (donc en particulier Heun est d'ordre 2).

Méthodes multi-pas

$$y_{n+s} = \left[b_0 f(t_n, y_n) + b_1 f(t_{n+1}, y_{n+1}) + \dots + b_s f(t_{n+s}, y_{n+s})\right] h$$
$$(a_s = 1) \qquad - \left(a_0 y_n + a_1 y_{n+1} + \dots + a_{s-1} y_{n+s-1}\right)$$

EE comme méthode multi-pas : $u_{n+1} = u_n + h f(t_n, u_n)$

$$\sum a_s y_{n+s} = \sum b_k f_{n+k} \cdot h \quad \Big| \; s = 1$$

$u_{n+1} - u_n = h f_n$; $a_1 = 1$, $a_0 = -1$, $b_0 = 1$, $b_1 = 0$

$$(1)\cdot y_{n+1} + (-1) y_n = 1 \cdot h f(t_n, y_n) + 0 \cdot h f(t_{n+1}, y_{n+1})$$

Demonstration du Thm 2.18 (consistance multi-pas)

Def Erreur de troncature d'un schema multi-pas est défini par

$$\tau_{n+s}(h) = \frac{\sum_{k=0}^{s} a_k X(t_{n+k}) - h \sum_{k=0}^{s} b_k f(t_{n+k}, X_{n+k})}{h}$$

Pour la consistance il faut $\tau(h) \to 0$ ($\tau_n^{(h)} \to 0 \ \forall n$).

$$X(t_{n+k}) = X(t_n) + \underbrace{(t_{n+k} - t_n)}_{\leq kh} \cdot X'(t_n) + \underbrace{O((t_{n+k} - t_n)^2)}_{O(h^2)}$$

$$f(t_{n+k}, X_{n+k}) = f(t_n, X_n) + \partial_t f(t_n, X_n) \cdot kh + \partial_x f(t_n, X_n) \cdot \underbrace{(X_{n+k} - X_n)}_{O(h)} + O(h^2) = f(t_n, X_n) + O(h)$$

$$\tau_{n+s}(h) = \frac{1}{h}\left[\sum_{k=0}^{s} a_k (X_n + kh X'_n)\right] - \sum_{k=0}^{s} b_k \underbrace{(f_n + O(h))}_{X'_n}$$

$$= \frac{1}{h}\left(\sum_{k=0}^{s} a_k\right) X_n + X'_n \left(\sum_{k=0}^{s} a_k \cdot k - \sum_{k=0}^{s} b_k\right) + O(h)$$

Pour consistance il faut que les termes en $\frac{1}{h}$ et h^0 soient nuls c'est à dire $\sum_{k=0}^{s} a_k = 0$ et $\sum_{k=0}^{s} k \cdot a_k = \sum_{k=0}^{s} b_k$.

$f = o(1) \sim O(h) \qquad O(\sqrt{h}) \quad O(h^0) \quad O(h^1) \ldots$

$O(h^k) \ \underline{k \gg 1}$

$\forall f_0 + \underbrace{f'(\cdot) \cdot h + \ldots}_{O(h)} + O(h^2) \ldots \qquad O(h) \subset o(1)$

EE $a_0 = -1, a_1 = 1, b_0 = 1, b_1 = 0 \quad \sum a_k = 0 \quad \sum k a_k = 1 = \sum b_k$

Adam-Bashford $s = 2 \quad a_2 = 1, a_1 = -1, a_0 = 0 \quad b_2 = 0, b_1 = \frac{3}{2}, b_0 = -1/2$

$\sum a_k = 0 \qquad \sum k a_k = 2 - 1 = 1 = \sum b_k$ ✓

BDF : $\nearrow D$

§3 Différentiation automatique, back propagation, optimisation, contrôle

(cf. cours en ligne www A2020)

But Calculer le gradient plr à un argument $x \mapsto \underbrace{f(x)}_{\in \mathbb{R}}$ ou une fonction $\beta(t) \to \underbrace{J[\beta(t)]}_{\in \mathbb{R}}$. $\beta(t)$ ici représenté par $\beta(t_n)$.

Ex 1 $f(x,y,z) = (3x^2+y)\cdot z$ $\quad \nabla f = \left(\frac{\partial f}{\partial x}(x_0,y_0,z_0), \frac{\partial f}{\partial y}(\dots), \frac{\partial f}{\partial z}(\dots) \right)$

$\partial_x f$
$\partial_1 f$

Ex 2 modèle SIR

$$\begin{cases} \frac{d}{dt} S(t) = -\beta(t) S(t) I(t) & t \in [0,T] \\ \frac{d}{dt} I(t) = \beta(t) S(t) I(t) - \gamma I(t) \end{cases}$$

But trouver le $\beta(\cdot)$ optimal (relié à l'effet de confinement, ...)

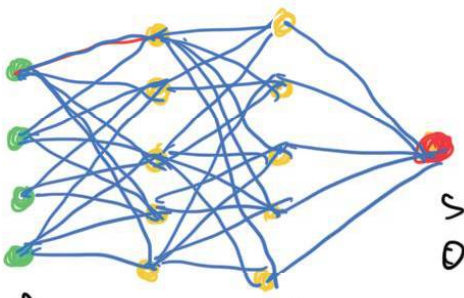
qui minimise $J[\beta] = \underbrace{S(0) - S(T)}_{\text{taille de l'épidemie}} + \underbrace{\int_0^T c(\beta(t))\,dt}_{\text{coût journalier du confinement (...)}}$

J dépend du $\beta(t)$ de 2 façons ① directement à travers $c(\beta)$

② indirectement car $S(T)$ dépend de $\beta(\cdot)$.

Pour optimiser il faut trouver $\nabla_\beta J[\beta]$.

Ex 3 Réseaux Neuronaux / Deep Learning / Neural Networks (NN)



sortie = fonction de l'entrée et des param des connexions

Pour entraîner le réseau il faut le gradient de la sortie plr aux param de connexion

Input X_0, Hidden 1 : X_1, Hidden 2 : X_2, Output X_3. (Vecteurs)

$X_1 = \underbrace{W_1}_{\text{matrice à optimiser}} X_0 + \underbrace{b_1}_{\text{vecteur param à optimiser}}$ $\quad X_0 \in \mathbb{R}^4 \quad W_1 \in \mathbb{R}^{5\times 4} \quad b_1 \in \mathbb{R}^5$ (colonne)

$X_1 \in \mathbb{R}^5 \quad X_2 \in \mathbb{R}^5 \quad X_3 \in \mathbb{R}^1$

$X_1^a = \mathrm{ReLU}(X_1) \qquad \mathrm{ReLU} : (\cdot)_+ \qquad X_2 = W_2 X_1^a + b_2$

$X_2^a = \mathrm{ReLU}(X_2)$

En pratique on optimise W_k, b_k et on doit calculer le gradient d'une fonction de la sortie X_3 p/r à W_k, b_k BEAUCOUP DE FOIS. En pratique il peut y avoir bcp de couches et dim des X_k grandes, ce qui donne bcp de paramètres $10^2 - 10^9$.

Ex d'utilisation X_0 = image $\quad X_3$ = décision chat / chien ($p \in [0,1]$)
↳ label 1 → label 0

Ex de fonctionnelle $\frac{1}{L}\left(\sum_{\ell=1}^{L} \left[F(X_3^{a\ell}) - \mathrm{Label}(X_0^\ell)\right]^2\right)$ fonction (cachée) des W_k, b_k.

Ex $W_1 = \begin{pmatrix} 1 & 2 & 1.2 & \sqrt{3} \\ 2 & 15 & -1 & 3 \\ \vdots & & & \end{pmatrix}$ $\qquad$ Ex F : sigmoïde $\frac{1}{1+e^{-x}}$

Pour optimiser il nous faut le gradient. Rq : pas de convexité, on utilise SGD. (Stochastic Gradient Descent) ou similaires.

1ère idée : utiliser une approximation différences finies $\Big\}$ on verra ceci servira à vérifier l'implémentation

$f(x+h) \overset{\text{Taylor}}{=} f(x) + f'(x)h + O(h^2) \Rightarrow \boxed{\frac{f(x+h)-f(x)}{h} = f'(x) + O(h)}$

(différence finie d'ordre 1, non centrée). $\boxed{f'(x) = \frac{f(x+h)-f(x)}{h} + O(h)}$

Autres formules (différences finies d'ordre 2)

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + O(h^2)$$ ← mieux que $O(h)$

Preuve (par Taylor) $\frac{f(x+h) - f(x-h)}{2h} = \frac{f(x) + f'(x)h + f''(x)\frac{h^2}{2} + f^{(3)}(\xi_1)\frac{h^3}{6} - \left(f(x) - hf'(x) + \frac{h^2}{2}f''(x) - \frac{h^3}{6}f^{(3)}(\xi_2)\right)}{2h} = f'(x) +$

$\frac{h^2}{12}\left(f^{(3)}(\xi_1) - f^{(3)}(\xi_2)\right) = f'(x) + O(h^2)$.

Utilisation Soit $G : \mathbb{R}^N \to \mathbb{R}$. Calculons $\nabla G(x)$

$$\frac{\partial G}{\partial x_k}(x) = \frac{G(x_1, \dots x_{k-1}, x_k+h, x_{k+1}, \dots, x_N) - G(x_1, \dots, x_{k-1}, x_k-h, \dots x_N)}{2h} + O(h^2)$$

(on applique la formule diff finies centrée à la fonction $\xi \to G(x_1, \dots, x_{k-1}, \xi, x_{k+1}, \dots, x_N)$. $f(\xi)$

On veut $f'(\xi)$.

Ex $N = \underbrace{5 \times 4 + 5}_{w_1 \quad b_1} + \dots$

Coût (en # d'évaluations de G) : pour calculer $\nabla_x G$ il faut $\geq N$ évaluations de G. Ceci est TROP GRAND lorsque N grand (ex 10^9).

OK si N petit (ex 1 $(\exists x^2+1)z - x$), ex 2 $\mathbb{R}$ si $\beta = \cos(t)$)

En général il faut utiliser d'autres méthodes (différenciation automatique) mais les diff finies serviront pour tester le code.

Attention à la précision $f^{\text{numérique}} \approx f^{\text{réel}} + O(10^{-16})$

$$f'_{\text{num}}(x) = \frac{f^{\text{réel}}(x+h) - f^{\text{réel}}(x-h) + O(10^{-16})}{2h} + O(h^2)$$

Donc $f'_{\text{numérique}}(x) = f'_{\text{réel}} + \underbrace{\frac{O(10^{-16})}{2h} + O(h^2)}_{O\left(\frac{10^{-16}}{h} + h^2\right)}$

Ex $h = 10^{-10}$ $\quad \frac{10^{-16}}{10^{-10}} + 10^{-20} \sim 10^{-6}$

10^{-3} : $\quad \frac{10^{-16}}{10^{-3}} + 10^{-6} \sim 10^{-6}$

Optimum : min plr à h : $\frac{10^{-16}}{h} + h^2$ sol $\frac{10^{-16}}{h} = h^2$

$h^3 \sim 10^{-16}$ $\boxed{h \sim 10^{-5,33}}$ Erreur $\simeq 10^{-10}$ pour $h = 10^{-6}$

Autos formules (ordre 3) : TD

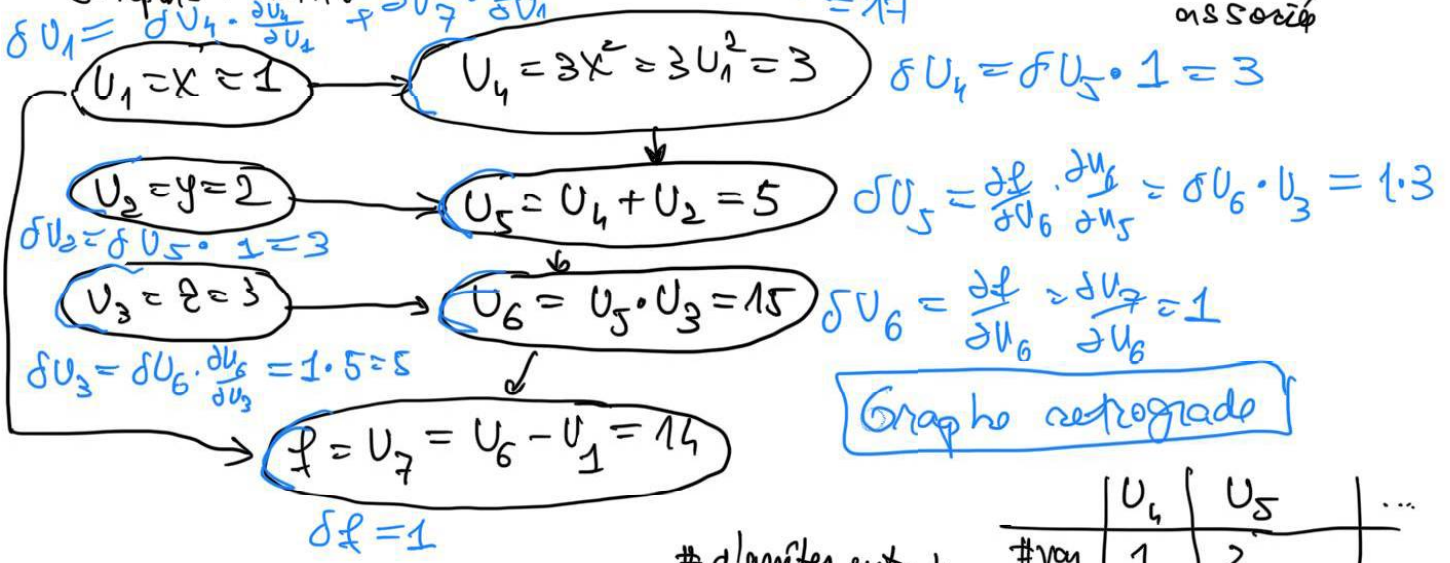
Chap : Diff Backward, ... M1 20/21

Graphe computationnel

$f(x,y,z) = (3x^2+y)\cdot z - x$

Mode direct / calcul de la sortie $f(\cdot)$: graphe computationnel direct.

Inputs : x, y, z sommets : valeurs + fonction associée



Graphe rétrograde

	U_4	U_5	...
#var	1	2	
fonct	$3U_1^2$	U_4+U_2	

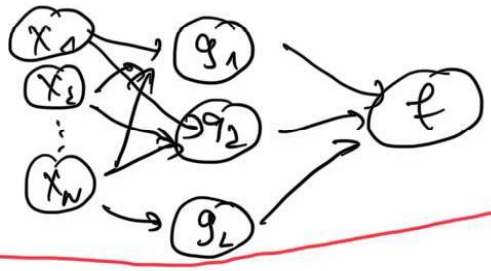
d'arêtes entrantes

Fonction du sommet # k : $\mathbb{R}^{\#} \to \mathbb{R}$

On veut $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$ ou encore $\frac{\partial f}{\partial U_1}, \frac{\partial f}{\partial U_2}, \frac{\partial f}{\partial U_3}$.

Notation $\frac{\partial f}{\partial U_k} =: \delta U_k$ $\frac{\partial f}{\partial f} = 1$ $\boxed{\delta U_7 = 1}$

Ex de structure



$X = (x_1, \dots, x_N)$

$$\frac{\partial f}{\partial x_k} = \frac{\partial}{\partial x_k} f(g_1(x), g_2(x), \dots, g_L(x)) = \sum_{\ell=1}^{L} \frac{\partial f}{\partial g_\ell}\cdot\frac{\partial g_\ell}{\partial x_k}$$

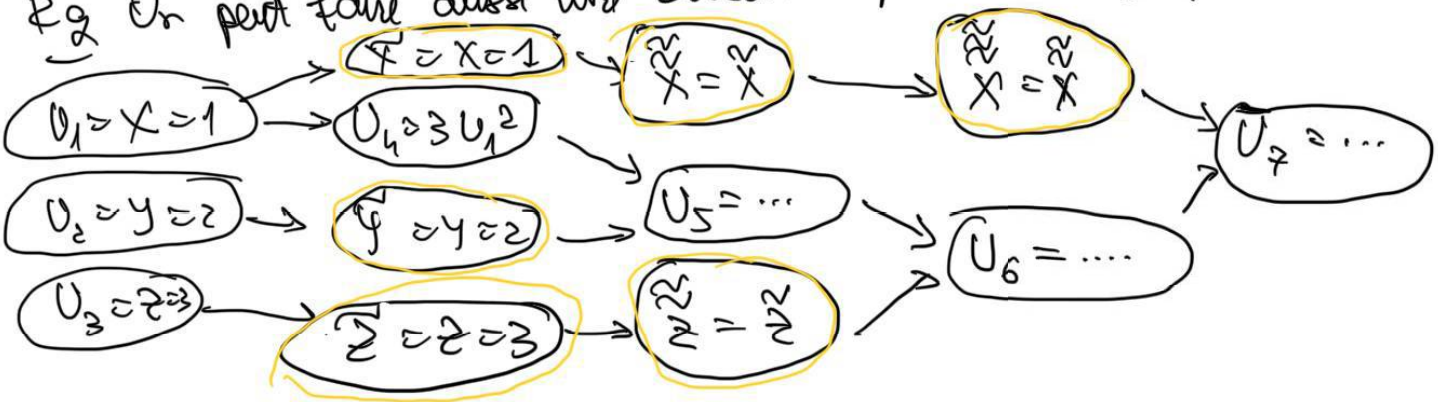
Calcul de $\frac{\partial f}{\partial g_\ell}$ possible dès le début ; il rend possible à la 2e étape le calcul de $\frac{\partial f}{\partial x_k}$. Une dérivée peut être calculée seulement après calcul des dérivées correspondant aux sommets reliés à ceux-ci par flèche sortante-entrante : ex δx_2 nécessite $\delta g_1, \dots \delta g_\ell$

$$\frac{\partial}{\partial x} f = \frac{\partial}{\partial x} f(U_1, U_4(\dots)) = \underbrace{\frac{\partial f}{\partial U_1}} \cdot \underbrace{\frac{\partial U_1}{\partial x}} + \underbrace{\frac{\partial f}{\partial U_4}}_{\delta U_4} \cdot \underbrace{\frac{\partial U_4}{\partial U_1}}$$

Prix / Coût de calcul des dérivées = coût d'un calcul direct = 1 évaluation de la fonction f.

Par contre les besoins en terme de mémoire sont plus grands.

Rq On peut faire aussi une version "plate" du graphe.



Exemple 1 Réseaux NN.

Exemple de tâche : base de données Ω d'exemples e.g. images $\omega \in \Omega$ = image. Fonction à minimiser : moyenne d'une fonction de perte ("loss" $\underset{\text{param à optimiser}}{x}, \underset{\text{exemple à moyenner}}{\omega} \longrightarrow L(x, \omega)$

On minimise $x \longmapsto \mathbb{E}_\omega L(x, \omega)$

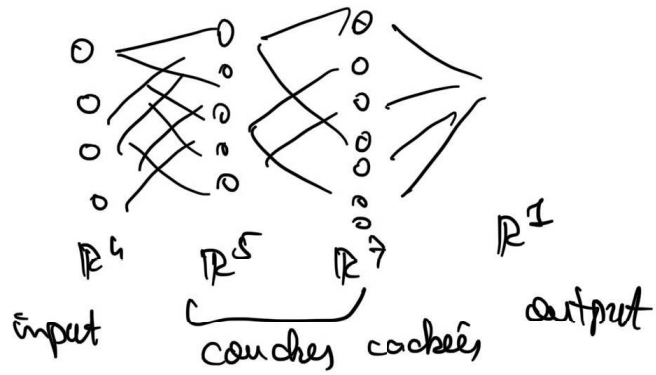
Ex de "loss" : ~~#~~ différence entre étiquette donnée par le NN et étiquette correcte.

Algo d'optimisation (ex) SGD "Stochastic Gradient Descent"

(ex d'autres Adam, Nesterov, momentum, SGD-G2, ...)

$x_{n+1} = x_n - \rho \nabla_x L(x_n, O_n)$ $\quad O_n$ ~~avec~~ i.i.d dans la base Ω avec une proba.

Ex de X



$y_0 \in \mathbb{R}^4$ $\quad \tilde{y}_1 = W_1 y_0 + b_1 \quad y_1 = \sigma_1(\tilde{y}_1)$ $\quad$ Ex $\sigma_1 = \mathrm{ReLU} = (\cdot)_+$

matrice $\mathbb{R}^{5\times 4}$ vecteur $\mathbb{R}^5$

$\tilde{y}_2 = W_2 y_1 + b_2 \quad y_2 = \sigma_2(\tilde{y}_2)$

$X = (W_1, b_1, W_2, b_2)$: à optimiser.

Exemple backward #1 : NN

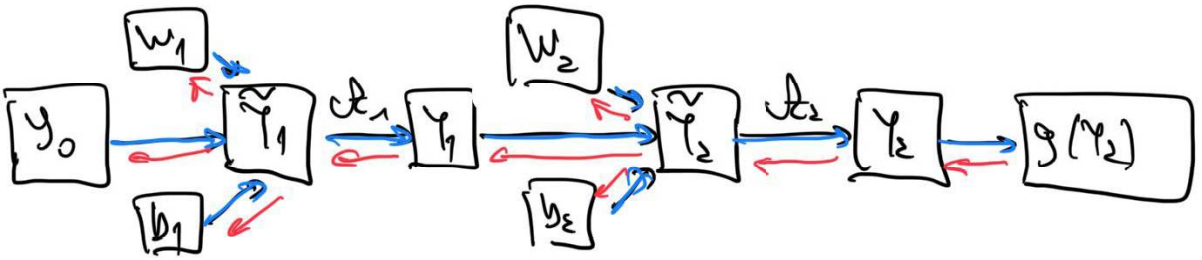
$\tilde{Y}_{k+1} = W_k Y_k + b_k$, $Y_{k+1} = \mathcal{A}_k(\tilde{Y}_{k+1})$ $\quad k = 0,1,2$

Rappel $Y_0 \in \mathbb{R}^4$, $Y_1 \in \mathbb{R}^5$, $Y_2 \in \mathbb{R}^7$, $Y_3 \in \mathbb{R}^2$ (output)

Ex $\mathcal{A}_1, \mathcal{A}_2 = ReLU = (\cdot)_+$, $\mathcal{A}_3$: exemple sigmoïde

$g = \mathcal{A}_3$ (sortie)

Graphe computationnel



Y_0 → $\mathcal{L}$(exemple, réseau) → W_k, b_k

Graphe direct : → backward ←

Avec les techniques vues précédemment on calcule les dérivées de $g(Y_2)$ p/r à W_k, b_k et donc les dérivées de la fonction loss p/r aux paramètres du réseau.

Ces dérivées sont par la suite utilisées pour optimiser le réseau avec, e.g., un algo SGD (ou Adam, ...).

à minimiser $X = (W_k, b_k)_{k=\dots}$ $\xrightarrow{\mathcal{L}}$ $E_w \mathcal{L}(X, w)$

↓ exemples

GD : $X_{n+1} = X_n - \gamma \nabla_X \mathcal{L}$

↓ pas d'apprentissage

→ coûteux à cause de la moyenne des E_w ...

$\nabla_X \mathcal{L} = E_w \nabla_X L(X, w)$

SGD : $X_{n+1} = X_n - \gamma \nabla_X L(X_n, w_n)$ w_n au hasard dans $\mathbb{R}$

Exemple #2 : contrôle

$x' = f(x(t), u(t))$ → contrôle

SIR $\begin{cases} \dot{S} = -\beta SI \\ \dot{I} = \beta SI - \gamma I \\ \dot{R} = \gamma I \end{cases}$ Etat $X = (S, I, R)$ $u(t) = \beta(t)$

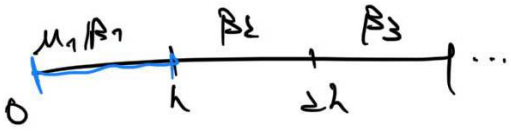
But minimiser $J(u) = \underbrace{S(0) - S(T)}_{\# \text{ malades}} + \int_0^T c(\beta(t))\,dt$

$= F(S(T))$ "taille de l'épidemie

Ex $c\left(\frac{1}{\lambda}\right) = \frac{1}{\lambda} \cdot c_0$ coût instantané du confinement

Pour optimiser $u(t)$ il faut calculer $\frac{\partial}{\partial u} J(u)$ (ensuite G.D. pour optimiser).

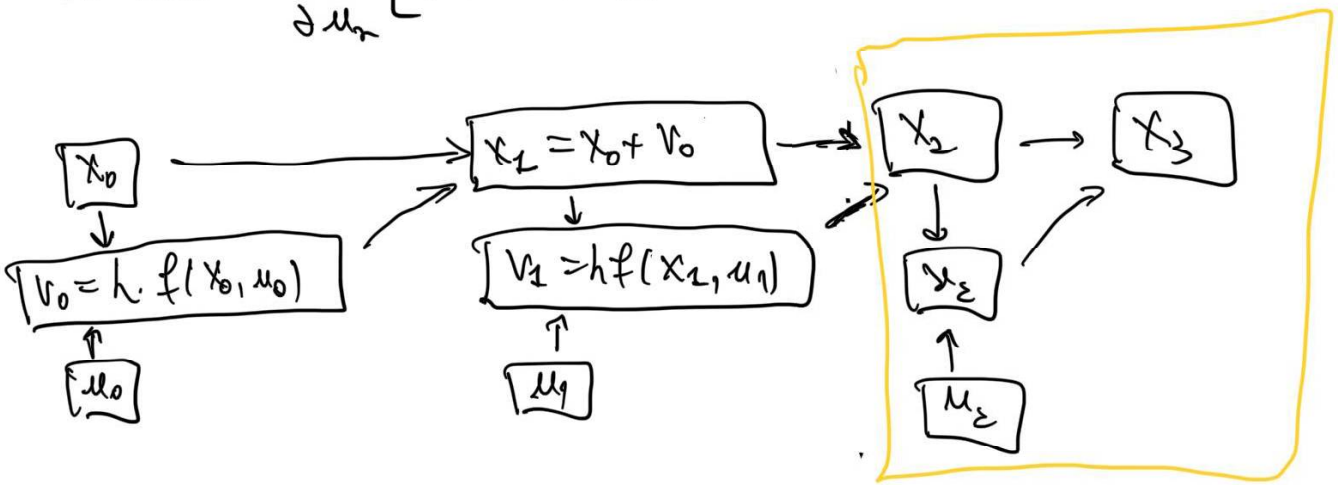
$$\frac{\partial}{\partial \beta(t)} \left[\int_0^T c(\beta(t))\,dt \right] = c'(\beta(t))$$



On veut aussi $\frac{\partial}{\partial \beta(t)} [S(T)] = ?$

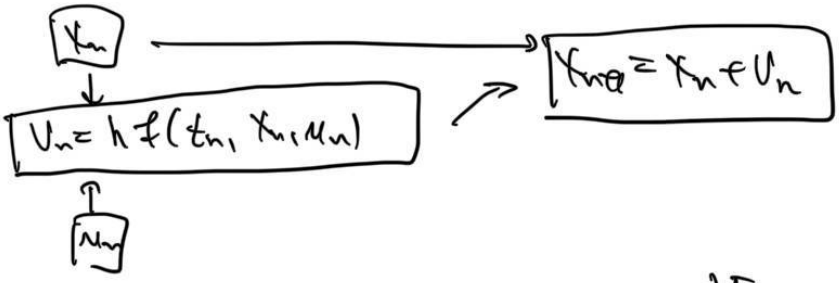
Supposons que pour le calcul de $S(T)$ on utilise EE. D'une manière générale $x_{n+1} = x_n + h f(x_n, u_n)$ EE

On veut $\frac{\partial}{\partial u_n} [F(x(T))]$.



$X'(t) = f(t, X(t), u(t))$ $\quad X_{n+1} = X_n + h f(t_n, X_n, u_n)$

cf. poly.



On note $\lambda_n = \delta X_n \left(= \frac{\partial F}{\partial X_n}\right)$ $\quad \lambda_{n+1} = \frac{\partial F}{\partial X_{n+1}}$

$\lambda_n = \frac{\partial F}{\partial X_n} = \frac{\partial F}{\partial X_{n+1}} \cdot \frac{\partial X_{n+1}}{\partial X_n} + \frac{\partial F}{\partial V_n} \cdot \frac{\partial V_n}{\partial X_n}$

$= \lambda_{n+1} \cdot 1 + \lambda_{n+1} \cdot h \cdot f'_x(t_n, X_n, u_n)$

$\lambda_{n+1} = \frac{\partial F}{\partial X_{n+1}}$

$\frac{\partial F}{\partial V_n} = \frac{\partial F}{\partial X_{n+1}} \cdot \frac{\partial X_{n+1}}{\partial V_n} = \lambda_{n+1} \cdot 1$

$\frac{\partial F}{\partial u_n} = \frac{\partial F}{\partial V_n} \cdot \frac{\partial V_n}{\partial u_n} = \lambda_{n+1} \cdot h \cdot f'_u(t_n, X_n, u_n)$

$\lambda_n = \lambda_{n+1} + h \lambda_{n+1} f'_x(t_n, X_n, u_n)$ $\longrightarrow$ c'est un schéma numérique

[ES] rétrograde par l'équation $\boxed{\lambda'(t) = -\lambda(t) \frac{\partial f}{\partial x}(t, X(t), u(t))}$

λ est dit : état adjoint.

$\min F(X(T))$: c'est une minimisation sous contrainte

$X' = f(t, X, u)$ $\longrightarrow$ Euler-Lagrange

$\min_{b(x) \le 0} a(x) \longrightarrow a(x) + \lambda b(x)$ $\quad (x, \lambda)$

Dans notre cas $F(X(T)) + \int_0^T [x' - f(t, X, u)] \lambda(t) dt =: G(X, \lambda)$

$\frac{\partial G}{\partial \lambda}$: on retrouve la contrainte ✓ $\quad \frac{\partial G}{\partial X(t)} = \frac{\partial}{\partial X(t)} \left[\int_0^T \lambda (x' - f(\dots))\right] =$

$\frac{\partial}{\partial X(t)} \left[\lambda(T) X(T) - \lambda(0) X(0) - \int \lambda' X - \int \lambda f(t, X, u) \right]$

$= \lambda' - \lambda f'_x(t, X, u)$ qui doit être nulle.

$\boxed{\frac{\partial}{\partial u(t)} = \lambda(t) \cdot \frac{\partial f}{\partial u}(t, X(t), u(t)) \cdot}$

Chapitre EDS

EDO $X'(t) = f(t, X(t))$ $\quad$ f connue $\quad \frac{dX(t)}{dt} = f(t, X(t))$

$dX(t) = f(t, X(t))\,dt$

Question comment traiter les aléas dans f, ou ailleurs ?

$dX(t) = f(t, X(t))\,dt$ + aléa $d\underbrace{W_t}_{\to \text{ mvt Brownien.}}$

-- Mvt Brownien ✓, Var quadratique ✓, Int Ito ✓, Formule Ito X

$[0, t]$ $\quad \frac{t}{N} = h \quad t_k = k \cdot h = \frac{k}{N} t$ $\quad t_0 \; t_1 \quad t = t_N$

Var quadratique $V_t^{(2)}(W, \cdot) = \sum_{k=0}^{N-1} \underbrace{(W_{t_{k+1}} - W_{t_k})^2}_{\sim \mathcal{N}(0,h)^2 \sim h\,\mathcal{N}(0,1)^2}$

$V_t^2(W, \cdot) \sim h\,\chi^2(N) \sim \underbrace{\frac{t}{N}\chi^2(N)} \sim t\left(\sum \frac{X_k^2}{N}\right) \overset{N\to\infty}{\sim} t$

$X_k \sim \mathcal{N}(0,1)$ iid $\qquad \downarrow \text{TCL}$ $\quad 1$

$\int f(x)\,d\underbrace{g(x)}$ $\quad$ R-S $\quad$ Portefeuille actif S_t, quantité θ_t

gain $t, t+\Delta t$: $\theta_t \cdot (S_{t+\Delta t} - S_t)$

$\lim_{\Delta t \to 0} \sum \text{gain}(t, t+\Delta t) = ?$

En gén on veut donner un sens à $\lim_{\Delta t \to 0} \sum_t H_t \cdot \underbrace{(W_{t+\Delta t} - W_t)}_{\text{"}dW_t\text{"}}$

On veut $\int H_t\,dW_t$ $\qquad \int 0\,dW_t = 0$

$\int 1\,dW_t = W_t - W_0$

$\int (H_1 + H_2)\,dW_t = \int H_1\,dW_t + \int H_2\,dW_t$

Ex $\int_0^T 3\,\underbrace{dW_t} = 3(W_T - W_0)$

$\int_0^t B_s\,dB_s = \lim_{|\Delta| \to 0} \sum_k B_{t_k}(B_{t_{k+1}} - B_{t_k}) = \frac{\left|\sum_k -(B_{t_{k+1}} - B_{t_k})\right|^2}{2}$

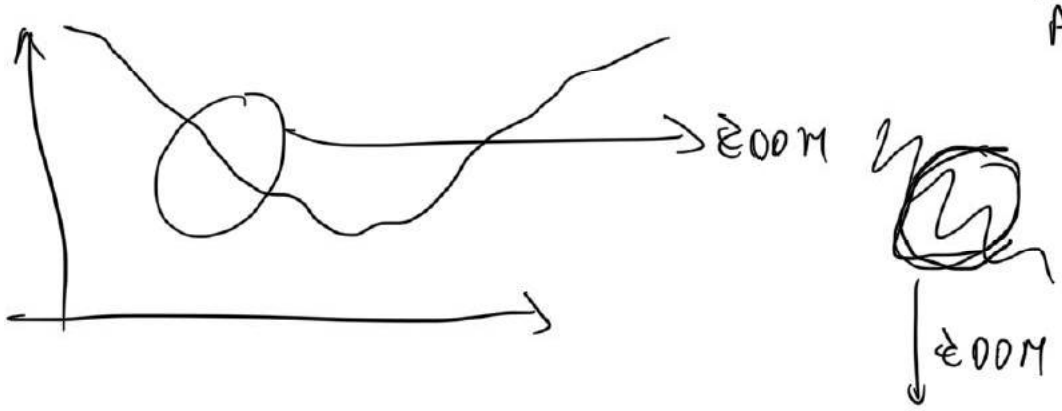
$$+B_{t_{k+1}}^2 - B_{t_k}^2 = \underbrace{\frac{1}{2}\sum -(B_{t_{k+1}} - B_{t_k})^2}_{\to -t/2} + \underbrace{\frac{1}{2}\sum_k \left(B_{t_{k+1}}^2 - B_{t_k}^2\right)}_{B_T^2 - B_0^2}$$

Donc $\int_0^T B_t \, dB_t = \frac{B_T^2 - T}{2}$. $\quad \int x\,dx = \frac{x^2}{2}$

$$\int_0^T 2 B_t \, dB_t = B_T^2 - T$$

4.1 item 2 : $\sum_k B_{\frac{t_k + t_{k+1}}{2}} (B_{t_{k+1}} - B_{t_k}) \longrightarrow ?$

$dX_t = \alpha_t dt + H_t dB_t \qquad \frac{dX_t}{dt} = \alpha_t + H_t \frac{dB_t}{dt}$ pb !



$\Delta t = 100 \quad \sqrt{\Delta t} = 10$ bruit petit par à la tendance

$\Delta t = 10^{-6} \quad \sqrt{\Delta t} = 10^{-3}$ bruit grand ...

$f(t, x) = x^2$. $\quad dX_t = \alpha_t dt + H_t dW_t \qquad Y_t = X_t^2$

$$df(t, X_t) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} H_t^2 dt \qquad \exists X_0$$

$\frac{\partial f}{\partial x} = 2x$

$\frac{\partial^2 f}{\partial x^2} = 2$

$$dY_t = 0 \cdot dt + 2 \cdot X_t \, dX_t + \frac{1}{2} \cdot 2 \cdot H_t^2 dt$$

$$dY_t = 2 X_t \, dX_t + H_t^2 dt$$

Ex $X_t = W_t \quad (\alpha = 0, H = 1)$

$$d(W_t^2) = 2 W_t \, dW_t + dt \, .$$

Exemples d'EDS $\boxed{dS_t = \mu S_t dt + \sigma S_t dW_t}$ σ, μ cst

modèle de Black-Scholes $S(0) = S_0$ donné

$a(t,x) = \mu \cdot x$ $b(t,x) = \sigma \cdot x$ OK

Montrons que $S_t = e^{(\mu - \sigma^2/2)t + \sigma W_t} \cdot S_0$

En $\sigma = 0$ $S' = \mu S$ $S_t = e^{\mu t} S_0$.

$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t = d(\mu t + \sigma W_t)$

$S_t = e^{\mu t + \sigma W_t} S_0$? NON

$dS_t = ?$ $f(t,x) = \left[e^{(\mu - \sigma^2/2)t + \sigma x} \right] \cdot S_0$

notation Y_t $\frac{\partial f}{\partial t} = \left(\mu - \frac{\sigma^2}{2}\right) f$ $\frac{\partial f}{\partial x} = \sigma f$ $\frac{\partial^2 f}{\partial x^2} = \sigma^2 f$

$$df(t,W_t) = \frac{\partial f}{\partial t}(t,W_t)\,dt + \frac{\partial f}{\partial x} dW_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} \cdot 1^2 \cdot dt$$

$$= \left(\mu - \frac{\sigma^2}{2}\right) f\,dt + \sigma f\,dW_t + \frac{1}{2}\sigma^2 f \cdot 1 \cdot dt$$

$dW_t = 0 \cdot dt + 1 \cdot dW_t$, d, $\mu = 1$

Donc $df(t,W_t) = \mu f\,dt + \sigma f\,dW_t$

Donc $Y_t = f(t,W_t) = e^{(\mu - \sigma^2/2)t + \sigma W_t}$ est bien solution de l'EDS.

Autre Ex Vasicek $dr_t = \alpha(\beta - r_t)\,dt + \sigma dW_t$ ...

$$d\left[\underbrace{\int_0^t a_s ds + \int_0^t b_s dW_s}_{dZ_t}\right] = \underbrace{a_t dt + b_t dW_t}_{dZ_t}$$

Schémas numériques pour EDS

$$dX_t = a(t, X_t)\,dt + b(t, X_t)\,dW_t$$

$$X_t = X_0 + \int_0^t a(s, X_s)\,ds + \int_0^t b(s, X_s)\,dW_s$$

Notation $t_0 = 0 < t_1 < t_2 < \dots < t_N = T =$ tps final

On veut un Y_n qui approche $X(t_n)$.

Schéma d'Euler-Maruyama (cqui de EE)

$$\begin{cases} Y_{n+1} = Y_n + a(t_n, Y_n)(t_{n+1} - t_n) + b(t_n, Y_n)\left[W_{t_{n+1}} - W_{t_n}\right] \\ Y_0 = X(t_0) \end{cases}$$

→ schéma explicite

→ on considère $t_{k+1} - t_k = cst = h$.

$\Delta W_n = W_{t_{n+1}} - W_{t_n}$, $a_n = a(t_n, Y_n)$, $b_n = b(t_n, Y_n)$

$$\boxed{Y_{n+1} = Y_n + a_n h + b_n \Delta W_n} \quad (\in M) \quad \boxed{Y_0 = X(t_0)}$$

Rq a_n, b_n : sont des v.a $\mathcal{A}_{t_n}$ mesurables

$(\mathcal{A}_t)$ = filtration de (W_t) (et $X(t_0)$).

(engendré par)

Si $b \equiv 0$ c'est le schéma de EE pour $X_t' = a(t, X_t)$

Schéma de Milshtein (Milstein, Mil'stein)

$$\begin{cases} Y_{n+1} = Y_n + a_n h + b_n \Delta W_n + \frac{1}{2} A_n b_n \left[(\Delta W_n)^2 - h \right] \\ Y_0 = Y(t_0) \end{cases} \quad (M)$$

(Supposons $a(t,X) = a(X)$, $b(t,X) = b(X)$ b' = dérivée de b $b'_n = b'(Y_n)$

Ce schéma est explicite.

Et pour les schémas implicites ? $a = \alpha X$, $b = \beta X$

Essai : $Y_{n+1} = Y_n + \alpha Y_{n+1} \cdot h + \beta Y_{n+1} \cdot \Delta W_n$

$\Rightarrow Y_{n+1} = \dfrac{Y_n}{1 - \alpha h - \beta \Delta W_n}$ pas bon !

car $\mathbb{E}\left(\dfrac{1}{|\mathcal{N}(0,1)|}\right) = \infty$!

On peut par contre prendre $Y_{n+1} = Y_n + \alpha Y_{n+1} \cdot h + \beta Y_n \Delta W_n$

semi-implicite

Pour (M) les termes aléatoires sont toujours $b_n \Delta W_n$ et $\dfrac{b_n b_n'}{2} \left[(\Delta W_n)^2 - h \right]$ et on peut mettre implicite $a(t_{n+1}, Y_{n+1}) \cdot h$.

Pour RK général : bcp. de termes.

Propriétés des schémas : CONSISTENCE (forte et faible).

Rappel EDO $\tau_n(h)$ = erreur de troncature $\tau = \max_n \tau_n$

Consistence : $\tau(h) \to 0$ (1)

ie. $\tau(h) \to 0$ $(h \to 0)$.

Def (Consistence faible) Un schéma donnant $(Y_n)_{n \geq 1}$ est dit faiblement consistant si

$$(W1) \quad \lim_{h \to 0} \mathbb{E}\left[\left| \mathbb{E}\left[\frac{Y_{n+1} - Y_n}{h} \,\middle|\, \mathcal{A}_{t_n} \right] - a_n \right|^2 \right] = 0$$

$$(W2) \quad \lim_{h \to 0} \mathbb{E}\left[\left| \mathbb{E}\left[\frac{(Y_{n+1} - Y_n)^2}{h} \,\middle|\, \mathcal{A}_{t_n} \right] - b_n^2 \right|^2 \right] = 0$$

Théorème Les schémas (E-M) et (M) sont faiblement consistants dès que a, b, b' sont bornées (ou assez régulières ...).

Dém a) Schéma (E-M) Vérifions (W1)

$$\mathbb{E}\left[\frac{Y_{n+1} - Y_n}{h} \,\middle|\, \mathcal{A}_{t_n} \right] - a_n = \mathbb{E}\left[\frac{Y_n + a_n h + b_n \Delta W_n - Y_n}{h} \,\middle|\, \mathcal{A}_{t_n} \right] - a_n$$

$$= \mathbb{E}\left[a_n + b_n \frac{\Delta W_n}{h} \mid \mathcal{F}_{t_n}\right] - a_n = a_n + b_n \mathbb{E}\left[\frac{\Delta W_n}{h} \mid \mathcal{F}_{t_n}\right] - a_n$$

$a_n \in \mathcal{F}_{t_n}$; $b_n \in \mathcal{F}_{t_n}$; $\frac{\Delta W_n}{h} \perp\!\!\!\perp \mathcal{F}_{t_n}$

$$= \frac{b_n}{h} \mathbb{E}\left[\Delta W_n \mid \mathcal{F}_{t_n}\right] = \mathbb{E}[\Delta W_n] = 0$$

Vérifions (W_2) :

$$\mathbb{E}\left[\frac{(Y_{n+1} - Y_n)^2}{h} \mid \mathcal{F}_{t_n}\right] - b_n^2 = \mathbb{E}\left[\frac{(a_n h + b_n \Delta W_n)^2}{h} \mid \mathcal{F}_{t_n}\right]$$

$$- b_n^2 = \mathbb{E}\left[a_n^2 h + 2 a_n b_n \Delta W_n + b_n^2 \frac{(\Delta W_n)^2}{h} \mid \mathcal{F}_{t_n}\right] - b_n^2$$

$\frac{(\Delta W_n)^2}{h} \sim \mathcal{N}(0,1)^2$

$$= h a_n^2 + 2 a_n b_n \underbrace{\mathbb{E}[\Delta W_n]}_{0} + b_n^2 \underbrace{\frac{\mathbb{E}(\Delta W_n)^2}{h}}_{=1} - b_n^2$$

$$\Rightarrow h a_n^2 \xrightarrow[h \to 0]{L^2} 0$$

b) Schéma de Talay

$$(W_1) : \mathbb{E}\left[\frac{Y_{n+1} - Y_n}{h} \mid \mathcal{F}_{t_n}\right] - a_n = \frac{\mathbb{E}\left[a_n h + b_n \Delta W_n + \frac{b_n b_n'}{2}(\Delta W_n^2 - h) \mid \mathcal{F}_{t_n}\right]}{h} - a_n$$

$$= a_n + b_n \underbrace{\mathbb{E}[\Delta W_n]}_{0} + \frac{b_n b_n'}{2} \underbrace{\mathbb{E}[\Delta W_n^2 - h]}_{0} - a_n = 0.$$

Rappels : $X \perp\!\!\!\perp \mathcal{F}$ et $\mathbb{E}[X \mid \mathcal{F}] = \mathbb{E}[X]$; $\mathbb{E}[XY \mid \mathcal{F}] = Y\,\mathbb{E}[X]$ si $Y \in \mathcal{F}$

$$(W_2)\quad \mathbb{E}\left[\frac{(Y_{n+1}-Y_n)^2}{h}\,\middle|\,\mathcal{A}_{t_n}\right]-b_n^2$$

$$=\frac{1}{h}\,\mathbb{E}\left[\left(a_n h+b_n\Delta W_n+\frac{b_n b_n'}{2}(\Delta W_n^2-h)\right)^2\,\middle|\,\mathcal{A}_{t_n}\right]-b_n^2$$

$$=\mathbb{E}\Big[a_n^2\cdot h+\cancel{b_n^2\frac{(\Delta W_n)^2}{h}}+\frac{b_n^2 b_n'^2}{4}\frac{(\Delta W_n^2-h)^2}{h}+\cancel{2a_n b_n\Delta W_n}$$

$$+\cancel{2\frac{a_n b_n b_n'}{2}(\Delta W_n^2-h)}+\frac{b_n^2 b_n'}{2}\underbrace{\cancel{\Delta W_n(\Delta W_n^2-h)}}\,\Big|\,\mathcal{A}_{t_n}\Big]-\cancel{b_n^2}$$

$$\underbrace{\mathbb{E}\,\Delta W_n^2}_{0}-h\underbrace{\mathbb{E}\,\Delta W_n}_{0}=\mathbb{E}\big(\underbrace{\Delta W_n(\Delta W_n^2-h)}_{\perp\!\!\!\perp\,\mathcal{A}_{t_n}}\big)=0.$$

$$=a_n h+\frac{b_n^2 b_n'^2}{4}\underbrace{\frac{\mathbb{E}(\Delta W_n^2-h)^2}{h}}=O(h)\xrightarrow[h\to 0]{L^2}0$$

car $\dfrac{(\Delta W_n^2-h)^2}{h}\sim\dfrac{(h\mathcal{N}^2(0,1)-h)^2}{h}=h\big(\underbrace{\mathcal{N}^2(0,1)}_{\chi^2(1)}-1\big)^2$

Rq (exo : La sol exacte vérifie (W_1) et (W_2) :

$$(W_1)\quad \mathbb{E}\left[\frac{X_{t_{n+1}}-X_{t_n}}{h}\,\middle|\,\mathcal{A}_{t_n}\right]-a(t_n,X_{t_n})$$

$$=\mathbb{E}\left[\frac{\int_{t_n}^{t_{n+1}}a(s,X_s)\,ds+\int_{t_n}^{t_{n+1}}b(s,X_s)\,dW_s}{h}\,\middle|\,\mathcal{A}_{t_n}\right]-a(t_n,X_{t_n})$$

$= \dots$ (développt de Taylor pour $a, b \dots$)

Déf (consistance forte) Un schéma $(Y_n)_{n \geq 1}$ est dit fortement consistant si :

(F1) : la condition (W1) est satisfaite

(F2) $\lim_{h \to 0} \mathbb{E}\left[\frac{1}{h} \left| Y_{n+1} - Y_n - \mathbb{E}(Y_{n+1} - Y_n \mid \mathcal{A}_{t_n}) - b_n \Delta W_n \right|^2 \right]$
$= 0$.

Thm Les schémas (EM) et (M) sont fortement consistants dès que a, b, b' sont bornés (...).

Déma a) (EM) (F1) ✓ (déjà vu)

(F2) : $Y_{n+1} - Y_n - \mathbb{E}[Y_{n+1} - Y_n \mid \mathcal{A}_{t_n}] - b_n \Delta W_n =$

$a_n h + b_n \Delta W_n - a_n h - b_n \Delta W_n = 0$ ✓

b) Milshtein (F1) ✓ ok déjà.

(F2) : $Y_{n+1} - Y_n - \underbrace{\mathbb{E}[Y_{n+1} - Y_n \mid \mathcal{A}_{t_n}]} - b_n \Delta W_n = a_n h +$

$b_n \Delta W_n + \underbrace{\frac{b_n b_n'}{2}} (\Delta W_n^2 - h) - \underbrace{a_n h} - b_n \Delta W_n$

$= \frac{b_n b_n'}{2} (\Delta W_n^2 - h)$. Il faut montrer que

$\lim_{h \to 0} \mathbb{E}\left[\frac{1}{h} \frac{b_n^2 b_n'^2}{4} (\Delta W_n^2 - h)^2 \right] = 0$ (=)

$\lim_{h\to 0} h \cdot \mathbb{E}\left[\frac{b_n^2 b_n'^2}{4}\left[\left(\frac{\Delta W_n}{\sqrt{h}}\right)^2 - 1\right]^2\right] = 0$ mais

car $b_n^2 b_n'^2$ borné et $\mathbb{E}\left[\left(\left(\frac{\Delta W_n}{\sqrt{h}}\right)^2 - 1\right)^2\right]$

$\Rightarrow$ borné.

Rq Il est possible de montrer que la consistance forte implique la consistance faible (ss hyp a, b, b' bornées).

Exo à faire ...

CONVERGENCE DES SCHEMAS EDS

Déf (convergence forte) Un schema EDS qui construit (Y_n) avec pas h est dit fortement convergent à l'ordre $\gamma > 0$ au temps T s'il existe $h_0 > 0$ et $C > 0$ indép de N et h (ici $N \cdot h = T$) telles que : $\mathbb{E}(|X(T) - Y_N|) \leq C h^\gamma \quad \forall h \leq h_0$.

Déf (Cv. faible) Le schéma (Y_n) est dit faiblement cv. à l'ordre $\beta > 0$ s'il $\exists h_0 > 0$, $C > 0$ indép de N, h (avec $Nh = T$) telles que

$$\left| \mathbb{E}[g(X(T))] - \mathbb{E}[g(Y_N)] \right| \leq C h^\beta \quad \forall h \leq h_0$$
$$\forall g \in C_p^{2p+2}$$

ou C_p^{2p+2} sont les fonctions de classe $2p+2$ à croissance au plus polynomiale à l'infini.

Rappel f est dite à croissance au plus polynomiale s'il $\exists\, s \in \mathbb{N}$ tel que $\exists\, C_s \in \mathbb{R}_+$ $|f(x)| \leq C_s(|x|^s + 1)$ $\forall x \in \mathbb{R}$.

Ex $f(x) = \sin(x)$ ✓ $x^{1/2}$ ✓ $\log x$ ✓ e^x non

Thm CV des schémas EM (et M) le schéma EM converge fortement à l'ordre 1/2 et faiblement à l'ordre 1. le schéma de Milshtein cv. fortement et faiblement à l'ordre 1.

Dem admis $\varepsilon_t = \sup_{0 \leq s \leq t} \mathbb{E}|Y_{n_s} - X_s|^2$ $n_s = \lceil s/h \rceil$.

Rq EM : la forme $b_n \Delta W_n$ est incomplet et cause la "perte" de 1/2 ordre de cv forte.

Question D'où vient la forme M par rapport à EM ?

Rappels Taylor sous forme intégrale (version opérateur) $X'(t) = a(t, X(t))$ $X_t = X + \int_0^t a(s, X_s)\,ds$

f fonction $\frac{d}{dt} f(X_t) = \frac{\partial f}{\partial x}(X_t) \cdot a(t, X_t)$

$f'(X_t) = f_x \cdot a$ $\quad L = a \cdot \frac{\partial}{\partial x}$ $\quad Lf = a \cdot f_x$

$$f(X_t) = f(X_0) + \int_0^t \underbrace{(Lf)(X_s)}_{\frac{d}{ds} f(X_s)}\,ds$$

$f(x) = x$ $\quad X_t = X_0 + \int_0^t a(X_s)\,ds$

$LX_s = a \cdot \frac{\partial}{\partial x} X = a$

Simplification $a = a(x)$

$f = a$ $\quad a(X_s) = a(X_0) + \int_0^s La(X_{s_2})\,ds_2$

$$X_t = X_0 + \int_0^t a(X_0)\,ds = X_0 + \int_0^t \left[a(X_0) + \int_0^s La(X_{s_2})\,ds_2 \right] ds$$

$$= X_0 + t \cdot a(X_0) + \int_0^t \int_0^s La(X_{s_2})\,ds_2\,ds .$$

Par récurrence on obtient Taylor intégrale :

$$f(X_t) = f(X_0) + \int_0^t Lf(X_s)\,ds = \dots$$

$$= f(X_0) + \sum_{k=1}^{n} \frac{t^k}{k!} (L^k f)(X_0) + \int_0^t \int_0^{s_1} \dots \int_0^{s_n} L^{n+1} f(X_{s_{n+1}})\, ds_{n+1} \dots ds_1$$

Taylor usuel est obtenu pour $a \equiv 1 \quad x' = 1$

$X_t = x_0 + t \quad t = x - x_0 \quad \mathcal{L} = a \frac{\partial}{\partial x} = \frac{\partial}{\partial x}$

Formules d'Ito - Taylor

$dX_t = a(X_t)dt + b(X_t)dW_t$

Ito $f(X_t) = f(X_0) + \int_0^t a(X_s) \frac{\partial}{\partial x} f(X_s) + \frac{1}{2} b(X_s)^2 \frac{\partial^2}{\partial x^2} f \Big\} ds$

$+ \int_0^t b \frac{\partial}{\partial x} f(X_s) dW_s .$

On évite b à la place de $b(X_{\dots})$ $f_x = \frac{\partial}{\partial x} f$

$f_{xx} = \frac{\partial^2}{\partial x^2} f \dots$

$$f(X_t) = f(X_0) + \int_0^t \mathcal{L}^0 f(X_s) ds + \int_0^t \mathcal{L}^1 f(X_s) dW_s$$

$\mathcal{L}^0 = a \frac{\partial}{\partial x} + \frac{1}{2} b^2 \frac{\partial^2}{\partial x^2} \qquad \mathcal{L}^1 = b \frac{\partial}{\partial x}$

$X_t = X_0 + \int_0^t a(X_s) ds + \int_0^t b(X_s) dW_s$

$= X_0 + \int_0^t \Big\{ a(X_0) + \int_0^s \mathcal{L}^0 a(X_\tau) d\tau + \int_0^s \mathcal{L}^1 a(X_\tau) dW_\tau \Big\} ds$

$+ \int_0^t \Big\{ b(X_0) + \int_0^s \mathcal{L}^0 b(X_\tau) d\tau + \int_0^s \mathcal{L}^1 b(X_\tau) dW_\tau \Big\} dW_s$

$= X_0 + a(X_0) \underbrace{\int_0^t ds}_{O(t)} + b(X_0) \underbrace{\int_0^t dW_s}_{O(t^{1/2})} + \text{reste}(R)$

$= X_0 + t \cdot a(X_0) + (W_t - W_0) b(X_0) + \text{reste}(R)$

$$R = \underbrace{\int_0^t \int_0^s L^0 a(X_\sigma)\, d\sigma ds}_{O(t^1)} + \underbrace{\int_0^t \int_0^s L^1 a(X_\sigma)\, dW_\sigma ds}_{O(t^{3/2})}$$

$$+ \underbrace{\int_0^t \int_0^s L^0 b(X_\sigma)\, d\sigma dW_s}_{O(t^{3/2})} + \underbrace{\int_0^t \int_0^s L^1 b(X_\sigma)\, dW_\sigma dW_s}_{O(t)}$$

$$L^1 b(X_\sigma) = L^1 b(X_0) + \underbrace{\int_0^\sigma L^0 L^1 b(X_u)\, du + \int_0^\sigma L^1 L^1 b(X_u)\, dW_u}_{O(t^{1/2})}$$

Donc

$$X_t = X_0 + a(X_0) \int_0^t ds + b(X_0) \int_0^t dW_s +$$

$$+ L^1 b(X_0) \int_0^t \int_0^s dW_\sigma dW_s + \text{reste}(R_2)$$

$$R_2 = O(t^{3/2})$$

$$R_2 = \int_0^t \int_0^s L^0 a(X_\sigma)\, d\sigma ds + \int_0^t \int_0^s L^1 a(X_\sigma)\, dW_\sigma ds + \iint_{0\,0}^{t\,s} L^0 b(X_\sigma)\, d\sigma\, dW_s$$

$$+ \int_0^t \int_0^s \int_0^\sigma L^0 L^1 b(X_u)\, du\, dW_\sigma dW_s + \iiint_{0\,0\,0}^{t\,s\,\sigma} L^1 L^1 b(X_u)\, dW_u dW_\sigma dW_s$$

Ordres : $2,\ 3/2,\ 3/2,\ 2,\ 3/2$ $\qquad O(t^{3/2})$.

Construction des schémas numériques :

EM : $X_t = X_0 + t\, a(X_0) + b(X_0)(W_t - W_0) + \{$ on néglige les autres $\}$

$$X_{t_{n+1}} = X_{t_n} + (t_{n+1} - t_n)\, a(X_{t_n}) + b(X_{t_n})(W_{t_{n+1}} - W_{t_n})$$

Ce schéma néglige le terme d'ordre 1

$$\underbrace{L^1 b(X_{t_n})}\ \underbrace{\int_{t_n}^{t_{n+1}} \int_{t_n}^{s} dW_\sigma dW_s}_{\simeq\ ?}$$

Mais $L^1 b(X_{t_n}) = b b'$

$$\int_{t_n}^{t_{n+1}} \int_{t_n}^{s} dW_r dW_s = \int_{t_n}^{t_{n+1}} (W_s - W_{t_n})\, dW_s$$

$$= \int_{t_n}^{t_{n+1}} W_s dW_s - W_{t_n} \int_{t_n}^{t_{n+1}} dW_s = \frac{W_{t_{n+1}}^2 - W_{t_n}^2}{2} -$$

$$- \frac{t_{n+1} - t_n}{2} - W_{t_n}(W_{t_{n+1}} - W_{t_n}) \quad \left(\text{car } \int B_t dB_t = \frac{B_t^2}{2} - \frac{t}{2} \right)$$

$$= \frac{1}{2}\left[W_{t_{n+1}}^2 - W_{t_n}^2 - 2 W_{t_n} W_{t_{n+1}} - h \right] = \frac{1}{2}\left[(W_{t_{n+1}} - W_{t_n})^2 - h \right]$$

Donc $X_{t_{n+1}} = X_{t_n} + a(X_{t_n})(t_{n+1} - t_n) + b(X_{t_n})(W_{t_{n+1}} - W_{t_n})$

$$+ \frac{1}{2} b(X_{t_n}) b'(X_{t_n}) \left[(W_{t_{n+1}} - W_{t_n})^2 - (t_{n+1} - t_n) \right] +$$

$$O\left[(t_{n+1} - t_n)^{3/2} \right] \quad \text{ce qui donne le schéma (M).}$$

Applications des EDS à l'évaluation des produits dérivés

Cf cours de processus stochastiques ! Cadre : $(\Omega, \mathcal{F}, \mathbb{P}), W$

Soit un produit dérivé de pay-off (final) G v.a. $\mathcal{F}_T$ mesurable

Ex $G = (S_T - K)_+$ $\quad \frac{dS_t}{S_t} = \mu dt + \sigma dW_t$ $\quad W$ = Brownien ss $\mathbb{P}, \mathcal{F}$

Il y a des cas où G n'est pas fonction de S_T, ex. options asiatiques qui payent $\left(\frac{1}{T}\int_0^T S_t dt - K\right)_+$.

On sait que (cf. cours proc. sto) qu'il y a une mesure dite "risque neutre" $\mathbb{Q}^*$ équivalente à $\mathbb{P}$ et $(\tilde{W}_t)_{t\geq 0}$ mvt-Brownien sous $\mathbb{Q}$ tels que :

$$C_t = \mathbb{E}^{\mathbb{Q}^*}\left[e^{-r(T-t)} G \mid \mathcal{F}_t\right] \quad (*)$$

($e^{-r(T-t)}$: actualisation ; $\mathcal{F}_t$: information connue en t)

$(G = C_T)$.

Par ailleurs $\frac{dS_t}{S_t} = r dt + \sigma d\tilde{W}_t$.

En fait $e^{-rt} C_t = \mathbb{E}^{\mathbb{Q}^*}\{e^{-rT} C_T \mid \mathcal{F}_t\}$

(e^{-rt} : actualisation du pay du produit dérivé) $\quad \tilde{C}_t = e^{-rt} C_t$

$$\tilde{C}_t = \mathbb{E}^{\mathbb{Q}^*}\{\tilde{C}_T \mid \mathcal{F}_t\}$$ pdlp martingale de $\tilde{C}_t$.

Supposons qu'on ait déjà montré que G / C_T peut être répliqué par un portefeuille Π_t : θ_t parts de ss-jacent S_t

reste : ss-risque au taux r

On dit aussi que Π_t finance G / C_T. Par AOA $\Pi_t = C_t$.

Comme Π_t est auto-financé $\quad d\Pi_t = \theta_t dS_t + (\Pi_t - \theta_t S_t) r dt$

Donc $d\Pi_t = r\Pi_t dt + \underbrace{\theta_t dS_t - \theta_t S_t dt}_{\theta_t S_t\{(\mu - r)dt + \sigma dW_t\}}$

Soit $\tilde{\Pi}_t = e^{-rt}\Pi_t$ Par Ito $\quad d\tilde{\Pi}_t = e^{-rt} d\Pi_t - r e^{-rt}\Pi_t dt$

Donc $d\tilde{\Pi}_t = e^{-rt}(d\Pi_t - r\Pi_t dt) = e^{-rt}\theta_t \cdot S_t \cdot$
$[(\mu - r)dt + r dW_t]$.

Soit $\mathbb{Q}^*$ une mesure obtenue de $\mathbb{P}$ par cgt de mesure telles que $d\tilde{W}_t = \frac{\mu - r}{\sigma} dt + dW_t$ et $\tilde{W}_t$ = $\mathbb{Q}^*$ mvt

Brownien.

Alors $d\tilde{\Pi}_t = e^{-rt}\theta_t S_t d\tilde{W}_t = \theta_t \tilde{S}_t d\tilde{W}_t$,
donc $\tilde{\Pi}_t$ = $\mathbb{Q}^*$-martingale ce qui donne la formule (E) car $\tilde{C}_t = \tilde{\Pi}_t = e^{-rt}\Pi_t$.

Ceci était ss l'hypothèse d'existence d'un portefeuille de couverture. En fait on peut même montrer ceci. Par exemple soit maintenant $M_t = \mathbb{E}^{\mathbb{Q}^*}[e^{-rT}G \mid \mathcal{F}_t]$ avec $\mathbb{Q}^*$ comme avant. Bien sûr si $\mathbb{E}^{\mathbb{Q}^*}\{G^2\} < \infty$ alors $(M_t)_{t \geq 0}$ est une $\mathbb{Q}^*$ martingale et alors par le thm de représentation ("toute martingale est une intégrale d'Itô")
$M_t = \int_0^t H_s d\tilde{W}_s$. Ceci nous permet de définir $\Pi_t = e^{rt} M_t$ qui est un portefeuille qui financera G / C_T. $\Pi_T = e^{rT} e^{-rT} G = G$.
$e^{rt} M_t = e^{rt} \int_0^t H_s \left(\frac{\mu - r}{\sigma} ds + dW_s\right)$ donc $\theta_t = e^{rt} H_t$.

On peut montrer que ss certaines hypothèses $\theta_t = \frac{\partial C_t}{\partial s}(t, S_t)$ c'est le delta-hedging (D-hedging) qui nous dit que le portefeuille autre finance

1 dérivé C_t	est un portefeuille ss risque.
$-\frac{\partial C_t}{\partial s}$ ss jacent S_t	
reste ss risque	

c'est cette dernière remarque qui permet de faire de répliquer d'options et de couvrir contre les risques venant de la vente d'options (en théorie). c'est ce que les vendeurs d'options font.

Apdv numérique pour calculer C_t :

1) on simule un mvt Brownien $\tilde{W}_t$ (donne M scénarios).

2) on simule S_t qui satisfait $\frac{dS_t}{S_t} = r dt + \sigma d\tilde{W}_t$

on produit scénarios $S_t^1, \dots (S_t^M)_{t \geq 0}$

3) on calcule la moyenne empirique

$$\frac{1}{M}\left[\sum_{k=1}^{M} h\left((S_t^k)_{t\geq 0}\right) \cdot e^{-rT}\right]$$

↳ proba risque-neutre, pas la proba historique

Soit $G = h((S_t)_t)$

Rq pour la simulation de S_t on part de S_0 (c'est la que $\mathcal{F}_0$ intervient) et on simule soit avec formule exacte soit avec (EM), (M'), ...

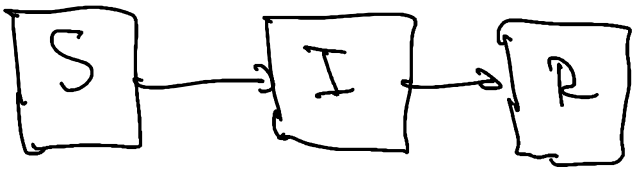
Ex $G = (S_T - K)_+$ Prix en 0 de l'option $= \frac{1}{M}\sum_{k=1}^{M} (S_T^k - K)_+ \cdot e^{-rT}$

Applications des EDO à la modélisation math en épidémiologie

Rappel SIR

$$\begin{cases} dS = -\beta SI/N \\ dI = \beta SI - \gamma I \\ dR = \gamma I \end{cases}$$

N = population totale $= S+I+R$



3 cases / états dans lesquelles peut se trouver une personne

S = "susceptible"

I = "infecté et infectieux"

R = "rétablis"

C'est le contact entre un individu infecté (classe I) et un susceptible (classe S) qui fait la maladie se propager.

On peut donc se dire que le nombre d'infectés nouveaux entre t et $t+\Delta t$ sera proportionnel à $S(t)$, $\frac{I(t)}{N}$, tx d'infection, Δt. On note β le facteur de proportionalité (β dép de tx de contact entre gens, tx d'infectiosité de la maladie ...)

Donc $S(t) - S(t+\Delta t) \simeq \beta S(t) \frac{I(t)}{N} \Delta t$

$$\Rightarrow \frac{S(t) - S(t+\Delta t)}{\Delta t} = \beta S(t) \frac{I(t)}{N}$$

$$\Rightarrow (\Delta t \to 0) \quad -S'(t) = \beta \frac{SI}{N} \qquad S' = -\beta S \frac{I}{N}$$

Pour le passage $I \to R$ on considère un tx de guérison constant et donc $R(t+\Delta t) - R(t) = \gamma \Delta t\, I(t) \xrightarrow{\Delta t \to 0}$

$R'(t) = \gamma I(t)$. L'ép de $I(t)$ découle de S', R'.

Propriétés du modèle SIR

$$\begin{cases} S' = -\beta \frac{SI}{N} & I' = \beta \frac{SI}{N} - \gamma I \\ R' = \gamma I \end{cases}$$

$\frac{1}{\gamma}$ = # d'unités de temps avant guérison / isolation

Question quand I est-il croissant ? $I' = \left(\beta \frac{S}{N} - \gamma\right) I$

$I \uparrow$ lorsque $I' > 0$ $\quad \frac{\beta S}{N} - \gamma > 0$ $\quad \frac{S}{N} > \frac{\gamma}{\beta}$

Si $S_0 \simeq N$ $\quad I' \simeq \gamma I \left(\frac{\beta}{\gamma} - 1\right)$ soit $R_0 = \frac{\beta}{\gamma}$

taux de reproduction

si $R_0 < 1$ $\quad I' \simeq \gamma I (R_0 - 1) < 0$: pas d'épidémie

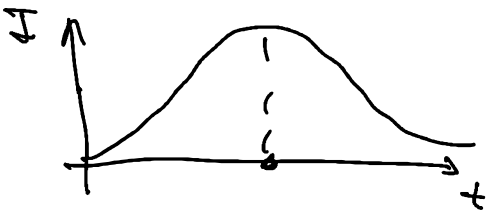
si $R_0 > 1$ $\quad I' \simeq \gamma I (R_0 - 1) > 0$: épidémie.

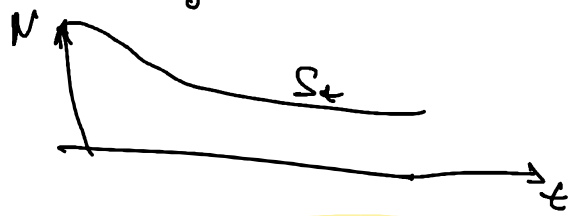
Si on peut diminuer R_0 pour le rendre < 1 : plus d'épidémie. Pour diminuer R_0 on peut soit diminuer β soit augmenter γ.

↓ diminuer les contacts

↓ quarantaine, isolation

$I' > 0$ lorsque $S_t \simeq N \cdot \frac{\gamma}{\beta} \simeq N \cdot \frac{1}{R_0}$ pic de l'épidémie.





Rq $S(\infty)$ n'est pas forcément 0 ! C'est l'immunité de groupe

$S(\infty)$ est solution d'une équation ...

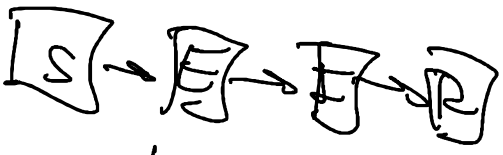
Taille de l'épidémie $I = S(0) - S(\infty)$.

$$1 - \frac{I}{S(0)} = e^{-R_0 (I + I_0)}$$

Rq Si $S_0 < N / R_0$ alors pas d'épidémie en particulier si on a vacciné $N - \frac{N}{R_0}$ individus.

Rq R_0 (covid-19) $\simeq 3,5$

Autres modèles SEIR



↓
infecté pas encore
infectieux "exposed"

SCIAR(D), ...

—

Calibration aux données pour trouver β, γ, ...

Examen : tout, 1 feuille A4 recto-verso admise.
Attention : regarder les corrections TD dans le poly mis à jour.