M1 MAPI3 - Simulations stochastiques 2024-2025
Xavier Bressaud *xavier.bressaud@math.univ-toulouse.fr*
Manon Costa *manon.costa@math.univ-toulouse.fr*
Bastien Mallein *bastien.mallein@math.univ-toulouse.fr*

TD 9-10 & TP 6 - Algorithme de Robbins-Monro

Exercice 1 -Regression et descente de gradient stochastique

Partie 1 - Régression linéaire : On suppose que l'on observe des couples (Y_i, X_i) i.i.d suivant un modèle de régression

$$Y_i = a^*X + b^* + \epsilon$$

où $\epsilon \sim \mathcal{N}(0,1)$.
On veut écrire une procédure de type Robbins-Monro pour estimer séquentiellement (a^*, b^*) à partir des données.

1) Vérifier que si $Y = a^*X+b^*+\epsilon$, le couple (a^*, b^*) est un minimum de la fonction $\phi(a,b) = \mathbb{E}[(Y-(aX+b))^2]$.

2) On rappelle que l'algorithme de Robbins-Monro pour trouver les points d'annulation d'un gradient $\nabla\phi(\theta) = \mathbb{E}(H(\theta, X))$ est donné par
$$\theta_{n+1} = \theta_n - \gamma_{n+1}H(\theta_n, X_{n+1})$$
pour (γ_n) une suite de pas positifs bien choisie et (X_n) une suite de variables aléatoires i.i.d. de même loi que X.
Proposer un algorithme de type Robbins-Monro pour estimer le couple (a, b) à partir des réalisations de (X_n, Y_n).

3) **TP -** On veut maintenant tester cet algorithme sur un jeu de donné. Charger le jeu de donné en utilisant les commandes :
```
U=np.loadtxt('TP4-exo2.txt')
X=U[:,0]
Y=U[:,1]
```
Afficher sur un graphique les données.

4) **TP -** Implémenter l'algorithme de Robbins Monro sur les données et tracer sur le graphique précédent la droite de regression obtenue ($y = a_nx + b_n$) (on utilisera successivement différentes suites de pas $\gamma_{n+1} = 0.2/(n+1)$, et $\gamma_{n+1} = 1.3/(n+1)$). Que pensez-vous de la convergence ?

5) **TP -** Comparer les valeurs obtenues avec l'estimateur des moindres carrés pour la régression linéaire. (Pour rappel, l'estimateur des moindres carré est donné par $\hat{a} = \frac{\sum(X_i-\bar{X})(Y_i-\bar{Y})}{\sum(X_i-\bar{X})^2}$ et $\hat{b} = \bar{Y} - \hat{a}\bar{X}$.)

Partie 2 - Modèle autoregressif : On considère ici le modèle autorégressif réel suivant :

$$X_{n+1} = \theta^*X_n + \epsilon_{n+1}$$

- La suite observée est $X = (X_n)_{n\geq 0}$ (on choisit $X_0 = 1$ arbitrairement).
- Le bruit $\varepsilon = (\varepsilon_n)_{n\geq 0}$ est une suite de variables aléatoires i.i.d. $\mathcal{N}(0, \sigma^2)$.
- Le paramètre $\theta^* \in (-1, 1)$ est inconnu.

On veut construire un algorithme de gradient stochastique pour estimer la valeur du paramètre θ^* inconnue. Pour cela, on suppose que l'on observe une suite de couple $(X_n, X_{n+1})_{n\geq 1}$ et que l'on souhaite minimiser la fonction
$$f(\theta) = \mathbb{E}((X_{n+1} - \theta X_n)^2).$$

1) Justifier que l'algorithme récursif
$$\theta_{n+1} = \theta_n + 2\gamma_{n+1}X_n(X_{n+1} - \theta_nX_n)$$
est l'algorithme de gradient stochastique associé.

2) **TP -** En choisissant une valeur de θ^* simuler un jeu de données correspondant au modèle auto-regressif.

3) **TP -** Coder l'agorithme de Robbins Monro pour ces données.
Étudier empiriquement la convergence presque sure de l'algorithme et le théorème de la limite centrale pour la suite de pas $\gamma_n = g_1/n$ et $g_1 < 1$.

Exercice 2 - Bandit à deux bras

Une machine à sous a deux leviers. La probabilité de gagner avec le levier A (respectivement B) est inconnue et vaut p_A (resp. p_B). Un joueur va chercher à optimiser son gain moyen, par un choix judicieux des leviers.

Voilà sa première stratégie : à l'instant $n+1$, il choisit le levier A si le taux de réussite du levier A (calculé sur les fois précédentes où le joueur a choisi le levier A) est supérieur ou égal au taux de réussite du levier B (calculé sur les fois précédentes où le joueur a choisi le levier B). Dans le cas contraire, il choisit le levier B. On initialise l'algorithme en disant que tant qu'on n'a pas joué avec un levier, le taux de réussite est nul.

1) On suppose que $p_A > p_B$. Expliquer pourquoi la situation suivante représente un piège susceptible d'arriver avec une probabilité non nulle : le joueur perd d'abord avec le levier A puis gagne avec le levier B.

Voilà sa seconde stratégie : à l'instant $n+1$, il choisit la machine A (resp. B) avec probabilité θ_n (resp. $1-\theta_n$) et

$$\theta_{n+1} = \begin{cases} \theta_n & \text{s'il a perdu} \\ \theta_n + \gamma_n(1-\theta_n) & \text{s'il a gagné avec } A \\ \theta_n - \gamma_n\theta_n & \text{s'il a gagné avec } B. \end{cases}$$

$0 < \theta_0 < 1$ est arbitraire et (γ_n) est une suite décroissante positive non sommable et de carré sommable (avec $0 < \gamma_0 < 1$).
Il est possible de montrer que θ_n converge vers 1 si A est la meilleure machine ($p_A > p_B$) et vers 0 sinon.

2) Montrer que $0 < \theta_n < 1$.

3) Calculer le champ moyen $\mathbb{E}[(\theta_{n+1} - \theta_n)/\gamma_n | \mathcal{F}_n]$ de l'algorithme.

4) Montrer que si $p_A > p_B$ alors $\mathbb{E}[\theta_n]$ est croissant .

5) Vérifier que si (A_n, B_n) sont des évènements indépendants de probabilités respectives, p_A, p_B et sont indépendants d'une suite de variables i.i.d uniformes (U_n) sur $[0,1]$ alors on peut écrire

$$\theta_{n+1} = \theta_n - \gamma_n \left(\theta_n \mathbf{1}_{B_{n+1}} \mathbf{1}_{U_{n+1} > \theta_n} - (1-\theta_n)\mathbf{1}_{A_{n+1}} \mathbf{1}_{U_{n+1} \le \theta_n})\right)$$

Pour quelle fonction h écrit-on un algorithme de Robbins-Monro ?
Vérifier que l'algorithme converge vers 0 ou 1 en fonction de la différence $p_A - p_B$.

7) **TP -** Illustrer cette stratégie par des simulations. En particulier, mettre en évidence la convergence du taux de réussite global vers $\max(p_A, p_B)$.

Exercice 3 - Estimation récursive d'un quantile

Soit (X_n) une suite de variables aléatoires i.i.d. de fonction de répartition F régulière et de densité $f > 0$. Pour $\alpha \in]0,1[$, on note θ_α le quantile d'ordre α de la loi F : $F(\theta_\alpha) = \alpha$. Afin d'estimer récursivement θ_α on utilise l'algorithme stochastique récursif :

$$\theta_{n+1} = \theta_n - \gamma_{n+1}(\mathbb{I}_{\{X_{n+1} \le \theta_n\}} - \alpha),$$

où θ_0 est arbitraire et $(\gamma_{n+1})_{n\ge 0}$ est une suite décroissante positive non sommable et de carré sommable.

1) Calculer le champ moyen $\mathbb{E}[(\theta_{n+1} - \theta_n)/\gamma_{n+1} | \mathcal{F}_n]$ de l'algorithme.

2) Utiliser le théorème de Robbins Monro pour montrer que l'algorithme converge vers θ_α. *On pourra prendre la fonction de Lyapunov* $V(\theta) = (\theta - \theta_\alpha)^2 + 1$

3) **TP -** On se place dans le cas où F est la fonction de répartition d'une loi normale standard $\mathcal{N}(0,1)$ et la suite de pas est $\gamma_{n+1} = 2/(n+1)$. Étudier numériquement la convergence de la suite (θ_n) dans le cas des trois quartiles $\alpha = 3/4$.

4) **TP -** Que pouvez-vous dire numériquement de la loi des fluctuations c'est à dire de la quantité $\sqrt{n}(\theta_n - \theta_\alpha)$. Que se passe-t-il pour une suite de pas $\gamma_{n+1} = 2(n+1)^{-2/3}$?

5) **TP -** Dans une situation de la vie réelle, le temps de calcul affecté à l'algorithme est limité. Aussi, on souhaite avoir des informations quantitatives sur l'erreur de l'algorithme à n fixé. Dans ce cas, on estime la vitesse de décroissance de l'erreur en norme L^2, c'est à dire de $\mathbb{E}(\theta_n - \theta_\alpha)^2))$ en fonction de n et de la suite de pas (γ_n) choisie.
Pour les deux suites de pas précédentes, représenter en fonction de n un estimateur Monte Carlo de $\mathbb{E}((\theta_n - \theta_\alpha)^2)$. Faire ensuite un diagramme log-log. Qu'en pensez vous ?