Consignes pour le travail demandé sur ce texte. On fera un exposé critique du modèle proposé. On rédigera en détails certaines affirmations du texte (voir aussi les suggestions de développements proposées à la fin du texte). Enfin, on proposera des idées d'illustrations informatiques en expliquant de façon algorithmique comment les mettre en oeuvre.
Durée : 3 h 30. Remettre ensuite votre copie dans le casier de S. Lemaire

---

Modélisation probabiliste pour l'optimistion des stratégies de maintenance de la voie ferrée[1]

La maintenance d'une voie ferrée vise à corriger les déformations liées au profil de la ligne, à la géologie, au climat, à l'âge et au trafic qu'elle supporte. On distingue quatre types de défauts de géométrie la voie : nivellement longitudinal, nivellement transversal, écartement, dressage. Ces déformations sont mesurés lors d'inspections effectuées à intervalles de temps réguliers. Pour des défauts légers, l'intervention la plus courante est le bourrage (correction de la géométrie de la voie par vibration et compactage mécanique du ballast). Des défauts trop importants nécessitent le remplacement du ballast, une intervention beaucoup plus coûteuse. On cherche donc à choisir l'intervalle entre les inspections et le seuil à partir duquel un bourrage est planifié de façon optimale pour réduire les coûts de maintenance.

1 Modélisation du système

On suppose que le vieillissement de la voie à un instant t est mesuré par une variable aléatoire positive X_t. En l'absence d'intervention, le processus $(X_t)_{t\geq 0}$ est croissant et évolue comme un processus Gamma de paramètres α et β (voir ci-dessous). Lors des interventions (bourrage ou remplacement du ballast), le processus $(X_t)_{t\geq 0}$ effectue un saut négatif correspondant au gain obtenu (amélioration de la géométrie de la voie). Notons $(X'_t)_{t\geq 0}$ le processus mesurant la dégradation de la voie sans tenir compte des interventions : le processus $(X'_t - X_t)_{t\geq 0}$ est donc constant entre deux interventions et effectue un saut positif lors des interventions. On suppose que le processus $(X'_t)_{t\geq 0}$ est indépendant de X_0 et est un processus Gamma de paramètres α et β, ce qui signifie que $X'_0 = 0$ et que pour tout $t, h \geq 0$, $X'_{t+h} - X'_t$ est indépendant de $(X'_s)_{0\leq s\leq t}$ et suit la loi Gamma de paramètres αh et β de densité donnée par

$$f_{X'_{t+h}-X'_t}(x) = \mathrm{I}_{[x>0]}\frac{\beta^{\alpha h}}{\Gamma(\alpha h)}x^{\alpha h-1}e^{-\beta x}.$$

On vérifie facilement que $\dfrac{X'_t}{t} \to \dfrac{\alpha}{\beta}$ p.s. quand $t \to +\infty$.
Les inspections de la voie sont effectuées à intervalles réguliers de durée $\triangle$. On observe donc les variables aléatoires $Y_n = X_{n\triangle}$. Il y a trois possibilités, suivant la position de Y_n par rapport à deux seuils $S > s > 0$.

– Si $Y_n < s$, il n'y a pas d'intervention pendant l'intervalle de temps $[n\triangle, (n+1)\triangle]$. Donc

$$Y_{n+1} - Y_n = X'_{(n+1)\triangle} - X'_{n\triangle}.$$

– Si $Y_n \geq s$, une intervention est effectuée avant l'instant $(n+1)\triangle$. L'intervention est un bourrage si $Y_n < S$ et un remplacement du ballast si $Y_n \geq S$. Dans ce cas, Y_{n+1} est à valeurs dans $[0, s[$ et est indépendante de tous les états antérieurs. On suppose que sa loi ne dépend pas de n et que son support contient 0. On note m son espérance.

[1]Texte remanié à partir d'un texte écrit par Christophe Leuridan pour la préparation à l'Agrégation de l'Université de Grenoble et inspiré de l'article de C. Meier-Hirmer, F. Sourget, M. Roussignol *Optimising the strategy of track maintenance*, Advances in Safety and Reliability (2005), 1385-1391.

2 Coût asymptotique

On peut montrer que la suite $(Y_n)_{n\in\mathbb{N}}$ définie par $Y_n = X_{n\triangle}$ est une chaîne de Markov récurrente (au sens où tout intervalle de $\mathbb{R}_+$ est visité une infinité de fois) et on admettra, par analogie avec le cas d'une chaîne de Markov à espace d'états fini qu'elle possède une unique probabilité invariante π. Le coût journalier moyen de la maintenance pendant l'intervalle de temps $[0, n\triangle[$ est

$$\frac{1}{n\triangle}\sum_{k=0}^{n-1}(C_i + C_b \mathrm{I}_{[s\leq Y_k<S]} + C_r \mathrm{I}_{[S\leq Y_k]})$$

où C_i, C_b, C_r sont les coûts respectifs d'une inspection, d'un bourrage et d'un remplacement de ballast. Quand $n \to +\infty$, on admettra, par analogie avec le cas d'une chaîne de Markov à espace d'états fini que ce coût journalier moyen tend vers

$$C_a = \frac{1}{\triangle}(C_i + C_b\pi([s, S[) + C_r\pi([S, +\infty[)).$$

Il suffit donc de connaître la probabilité invariante pour déterminer le coût asymptotique.

3 Un modèle simplifié

Dans cette partie, on étudie un modèle simplifié où $\alpha\triangle = 1$ et dans lequel s, S et les Y_n sont entiers. On suppose que les accroissements $X'_{(n+1)\triangle} - X'_{n\triangle}$ sont indépendants et de loi géométrique sur $\mathbb{N}^*$ de paramètre p. On a $Y_n = Y_0 + X'_{n\triangle}$ tant que $n \leq N_1$ où $N_1 = \inf\{n \in \mathbb{N} : Y_n \geq s\}$ est le numéro de la première inspection suivie d'une intervention.
Sachant que $Y_0 = y < s$, les variables aléatoires N_1 et $Y_{N_1} - s$ sont donc indépendantes : $N_1 - 1$ suit la loi binomiale $\mathcal{B}(s - y - 1, p)$ et $Y_{N_1} - s$ suit la loi géométrique sur $\mathbb{N}$ de paramètre p. En particulier,

$$\mathrm{E}[N_1|Y_0 = y] = 1 + p(s - y - 1) \text{ et } P[Y_{N_1} \geq S|Y_0 = y] = (1 - p)^{(S-s)}.$$

Sachant que $Y_0 = x \geq s$, la variable aléatoire Y_1 est strictement inférieure à s et d'espérance m ; le premier instant de retour R_s de la chaîne $(Y_n)_{n\in\mathbb{N}}$ dans $[s, +\infty[$ a pour espérance $1 + p(s - m - 1)$. On en déduit les valeurs de $\pi([s, +\infty[)$ et de $\pi([S, +\infty[)$ ce qui permet de calculer C_a.

Suggestions de développements

- *Partie 1 :*
  1. Si $t, h \geq 0$, quelle est la 1oi de X'_t, de $X'_{t+h} - X'_t$ et de X'_{t+h} ? Quelles propriétés des lois Gamma rendent compatibles entre elles les hypothèses faites sur le processus X' ?
  2. Justifier le fait que $\frac{X'_t}{t} \to \frac{\alpha}{\beta}$ p.s. quand $t \to +\infty$.
  3. Si $(Z_1, \ldots, Z_n)$ est un échantillon de la loi Gamma de paramètres $\alpha\triangle$ et β, quelle est la densité de $(Z_1, \ldots, Z_n)$? En déduire que l'estimateur du maximum de vraisemblance de (α, β) est donné par $\frac{\hat{\alpha}\triangle}{\hat{\beta}} = \frac{1}{n}(Z_1 + \ldots + Z_n)$ et $\psi(\hat{\alpha}\triangle) - \ln\hat{\beta} = \frac{1}{n}(\ln Z_1 + \ldots + \ln Z_n)$ en notant pour tout $s > 0$,

     $$\psi(s) = \frac{\Gamma'(s)}{\Gamma(s)} = \lim_{\ell\to+\infty} \ln\ell - \sum_{k=0}^{\ell}\frac{1}{s+k}.$$

     Que se passe-t-il quand $n \to +\infty$? (*Indication :* on pourra montrer que l'application $\psi - \ln$ est strictement croissante sur $\mathbb{R}_+^*$).
- *Partie 2 :*
  1. Justifier le fait que $(Y_n)_{n\in\mathbb{N}}$ soit une chaîne de Markov. Quelle est la loi de Y_{n+1} sachant que $Y_n = x$? Calculer $\mathrm{E}[Y_{n+1}|Y_n]$.
  2. Donner une interprétation de l'expression de C_a.
  3. Expliquer comment estimer C_a par simulations.
- *Partie 3 :* Justifier les calculs effectués dans le cas du modèle simplifié et donner les propriétés de la chaîne de Markov $(Y_n)_n$.