IFT 3245
Simulation et modèles

Fabian Bastin
DIRO
Université de Montréal

Automne 2016

Variables antithétiques (AV)

Semblable à l'idée des VAC, sauf que l'on veut maintenant estimer une seule espérance par un moyenne de plusieurs estimateurs négativement corrélés.

Supposons que l'on a k estimateurs sans biais de μ, $(X^{(1)}, \ldots, X^{(k)})$, avec $\mathrm{Var}[X^{(1)}] = \cdots = \mathrm{Var}[X^{(k)}]$.

La moyenne

$$X_{\mathrm{a}} = \frac{1}{k} \sum_{\ell=1}^{k} X^{(\ell)}$$

est un estimateur sans biais de μ, avec variance

$$\mathrm{Var}[X_{\mathrm{a}}] = \frac{1}{k^2} \sum_{j=1}^{k} \sum_{\ell=1}^{k} \mathrm{Cov}[X^{(j)}, X^{(\ell)}] = \frac{\mathrm{Var}[X_{(1)}]}{k} + \frac{2}{k^2} \sum_{j<\ell} \mathrm{Cov}[X^{(j)}, X^{(\ell)}]$$

Variables antithétiques

Si les $X^{(\ell)}$ sont indépendants, les covariances sont nulles.

La variance est réduite ssi $\mathrm{Var}[X_{\mathrm{a}}] < \mathrm{Var}[X_{(1)}]/k$,
ssi la somme des covariances est négative.

Pour un total de $n = mk$ répétitions, on définit

$$\bar{X}_{\mathrm{a,n}} = \frac{1}{m}\sum_{i=1}^{m} X_{\mathrm{a,i}} = \frac{1}{n}\sum_{i=1}^{m}\sum_{\ell=1}^{k} X_i^{(\ell)},$$

où $X_{\mathrm{a},1}, \ldots, X_{\mathrm{a,m}}$ sont des copies i.i.d. de X_{a}.

Estimateur sans biais de $\mathrm{Var}[X_{\mathrm{a}}]$; variance empirique:

$$S^2_{\mathrm{a,m}} = \frac{1}{m-1}\sum_{i=1}^{m}\left(X_{\mathrm{a,i}} - \bar{X}_{\mathrm{a,n}}\right)^2.$$

Méthode AV générale

Soit $\mu = E[X]$, où $X = f(\boldsymbol{U})$ et $\boldsymbol{U} = (U_1, U_2, \dots)$ une suite i.i.d. $U(0,1)$. Pour $\ell = 1, \dots, k$, soit $X^{(\ell)} = f(\boldsymbol{U}^{(\ell)})$ où $\boldsymbol{U}^{(\ell)}$ est une suite i.i.d. $U(0,1)$.

Le but de l'approche de variables antithétiques généralisées est d'induire une dépendance entre les $\boldsymbol{U}_\ell$ de manière à rendre $\sum_{j<\ell} \mathrm{Cov}[X^{(j)}, X^{(\ell)}] < 0$.

Paires antithétiques

On a $k = 2$,
$\boldsymbol{U}^{(1)} = \boldsymbol{U} = (U_1, U_2, \dots)$,
$\boldsymbol{U}^{(2)} = \mathbf{1} - \boldsymbol{U} = (1 - U_1, 1 - U_2, \dots)$, la suite antithétique.

Posons $X_{\mathrm{a}} = [X^{(1)} + X^{(2)}]/2$,

$$\begin{aligned} \mathrm{Var}[X_{\mathrm{a}}] &= (\mathrm{Var}[X^{(1)}] + \mathrm{Var}[X^{(2)}] + 2\,\mathrm{Cov}[X^{(1)}, X^{(2)}])/4 \\ &= (\mathrm{Var}[X^{(1)}] + \mathrm{Cov}[X^{(1)}, X^{(2)}])/2, \end{aligned}$$

et $\mathrm{Var}[X_{\mathrm{a}}] < \mathrm{Var}[X]/2$ ssi $\mathrm{Cov}[X^{(1)}, X^{(2)}] < 0$.

Intuition: Les événements désastreux pour $X^{(1)}$ seront compensés par des événements antithétiques chanceux pour $X^{(2)}$, et vice-versa.

Mise en oeuvre

Meilleur cas possible: pour une fonction linéaire la variance est réduite à zéro.
Pire cas: si $X^{(1)}$ et $X^{(2)}$ sont parfaitement corrélées, la variance est doublée (e.g., $f(\boldsymbol{U}) = |U_1 - 1/2|$).

Pour minimiser la covariance, il faudrait générer directement $X^{(1)} = F^{-1}(U)$ et $X^{(2)} = F^{-1}(1-U)$, où $U \sim U(0,1)$ et F est la fonction de répartition de X.
Pas pratique.

Comme pour les VAC, on appliquera l'inversion pour maximiser la covariance à un plus bas niveau.

Combinaison VAC-AV

Supposons que l'on veut comparer deux systèmes en utilisant des VAC, et aussi des paires de répétitions antithétiques.
Si ces deux méthodes fonctionnent bien séparément, fera-t-on nécessairement mieux en les combinant ? Non.
Soit $X_k^{(1)} = f_k(\boldsymbol{U})$ et $X_k^{(2)} = f_k(\mathbf{1} - \boldsymbol{U}))$, $k = 1, 2$. On a

$$\mathrm{Var}\left[\frac{X_2^{(1)} + X_2^{(2)}}{2} - \frac{X_1^{(1)} + X_1^{(2)}}{2}\right]$$

$$= \frac{\mathrm{Var}[X_1^{(1)}] + \mathrm{Var}[X_1^{(2)}] + \mathrm{Var}[X_2^{(1)}] + \mathrm{Var}[X_2^{(2)}]}{4}$$

$$+ \frac{\mathrm{Cov}[X_1^{(1)}, X_1^{(2)}] + \mathrm{Cov}[X_2^{(1)}, X_2^{(2)}]}{2} - \frac{\mathrm{Cov}[X_1^{(1)}, X_2^{(1)}] + \mathrm{Cov}[X_1^{(2)}, X_2^{(2)}]}{2}$$

$$- \frac{\mathrm{Cov}[X_1^{(2)}, X_2^{(1)}] + \mathrm{Cov}[X_1^{(1)}, X_2^{(2)}]}{2}.$$

$$= \mathrm{Var}[X_1^{(1)}] + \underbrace{\mathrm{Cov}[X_1^{(1)}, X_1^{(2)}]}_{\leq 0} - \underbrace{\mathrm{Cov}[X_1^{(1)}, X_2^{(1)}]}_{\geq 0} - \underbrace{\mathrm{Cov}[X_1^{(2)}, X_2^{(1)}]}_{?}.$$

Cas multidimensionnel

Un échantillonnage est AV-concordant si $U_{2,j} = 1 - U_{1,j}$ pour certaines coordonnées j telles que f est monotone par rapport à U_j, tandis que $U_{2,j}$ est indépendant de $U_{1,j}$ pour les autres coordonnées.

Theorème

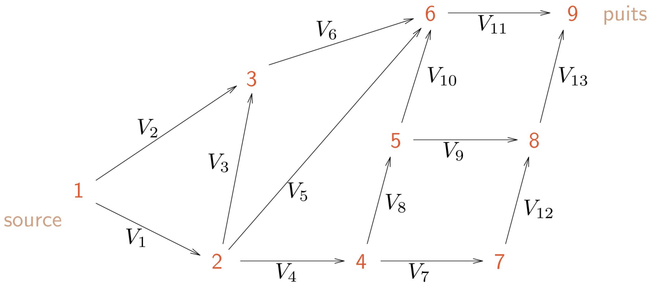
Pour un échantillonnage AV-concordant, on a $Cov[X^{(1)}, X^{(2)}] \leq 0$.

Un réseau d'activités stochastique.

Graphe acyclique $(\mathcal{N}, \mathcal{A})$, de noeuds $\mathcal{N}$ et arcs $\mathcal{A}$ (les activités). Le graphe donne les relations de précédence entre les activités. Chaque activité $k \in \mathcal{A}$ a une durée aléatoire V_k de fonction de répartition F_k.

Si V_k est la longueur de l'arc k, la durée minimale du projet est la longueur T du plus long chemin dans le réseau.

On peut vouloir estimer $P[T > x]$ pour un certain x, par exemple.



Exemple: réseau d'activités stochastique

On veut estimer $\mu = P[T > x]$, où T est la longueur du plus long chemin et x une constante.

On génère $V_j = F_j^{-1}(U_j)$ pour $X^{(1)}$ et $V_j = F_j^{-1}(1 - U_j)$ pour $X^{(2)}$.

Puisque T est non-décroissante en $V_j = F_j^{-1}(U_j)$, qui est non-décroissante en U_j, on a un échantillonnage AV-concordant et la variance est réduite.

En pratique, sur ce genre d'exemple, la variance est réduite environ de moitié.