algorithme de correction

Wilfried Ehounou

November 25‹ 2017

Contents

0.1 Correction de la matrice de corrélation

Si le graphe $G_C = (V_C, E_C)$ est un graphe de corrélation alors l'algorithme de couverture determine sa line couverture $\mathcal{C}$. En effet, par récurrence sur l'ensemble des sommets, on montre à chaque étape qu'il existe un sommet non encore couvert qui soit :

- est couvert par une clique appartenant à $\mathcal{C}$ et son voisinage restant et lui peuvent être converts par une nouvelle clique.
- n'est couvert par aucune clique de $\mathcal{C}$ et son voisinage restant et lui peuvent être couverts par une ou deux nouvelles cliques.

Dans le cas où la line-couverture de G_C ne peut être fournie à cause des erreurs de corrélations, nous avons des sommets couverts par soit aucune clique ou soit par plus de deux cliques. Ces sommets, labellisés à -1, forment l'ensemble $sommets\_1 = \{\exists z \in V, Cliq(z) = -1\}$ et sont appelés *sommets à corriger*.

Nous proposons l'*algorithme de correction* qui va modifier l'ensemble initial E_C par ajout et suppression d'arêtes dans le but d'obtenir un *line graphe*. Nous allons considérer un ordre $O_z = [z_1, z_2, \cdots, z_t]$ de sommets de $sommets\_1$ qui correspond au mode de sélection de ceux-ci pendant la phase de correction. Il en suit que l'ordre a une influence sur le line graphe fourni parce que la correction modifie le voisinage des sommets. Il est montré dans la chapitre **??**.
Soit E_C^i l'ensemble des arêtes de G_C après le traitement des $i-1$ premiers sommets dans l'ordre O_z. De même, on note $\mathcal{C}^i$ l'ensemble des cliques de G_C à l'étape i et donc $E_C^1 = E_C$ et $\mathcal{C} = \mathcal{C}^1$.

Soient $z = z_i$ le i-ième sommet et $\mathcal{C}(z) = \{C_1, \cdots, C_k\}$ l'ensemble des cliques de $\mathcal{C}^i$ de taille supérieure ou égale à 3 auxquelles le sommet z appartient. Notons que, par définition et par construction, chaque paire de cliques dans $\mathcal{C}(z)$ n'a que z comme sommet commun et que $S(z)$ est l'union des voisins v de z dans des cliques $\{v, z\} \in \mathcal{C}^i$ de taille 2 et des voisins v de z tels que l'arête $[z, v]$ n'est couverte par aucune clique de $\mathcal{C}^i$.

$$C(z) = \{C_i, i \in [1, k] \mid |C_i| \geq 3 \ \& \ C_i \in \mathcal{C}^i\} \tag{1}$$

$$S(z) = \{v \in \Gamma_G(z) \mid \{v, z\} \in \mathcal{C}^i\} \cup \{v \in \Gamma_G(z) \mid \nexists C \in \mathcal{C}^i, [z, v] \in E_C(C)\} \tag{2}$$

Définition 1 *Deux cliques C et C' de $\mathcal{C}(z)$ sont* contractables *si aucune arête $[u, v]$ de E_C^i telle que $u \in C$ et $v \in C'$ n'est couverte par une clique (autre que u, v) dans $\mathcal{C}$. Un ensemble de cliques de $\mathcal{C}$ est contractable si tous les cliques sont deux à deux contractables.*

Définition 2 *Une clique $C \in C_i$ est* voisine *de z si $C \notin \mathcal{C}(z)$ et $card(C \cap S(z)) \geq 2$. La dépendance d'une clique C voisine de z est l'ensemble $D_z(C) \subset \mathcal{C}(z)$ tel que $C' \in D_z(C)$ si et seulement si $C' \cap C \cap \Gamma_G(z) \neq \emptyset$.*
Une clique C est augmentante *pour le sommet z si et seulement si elle est voisine de z et $D_z(C)$ est vide ou $D_z(C) \cup \{C\}$ est contractable.*

$$voisine(z) = \{C \in \mathcal{C}^i \ \mid \ C \notin C(z) \ \& \ card(C \cap S(z)) \geq 2\} \tag{3}$$

$$D_z(C) = \{C' \in C(z) \ \mid \ C' \cap C \cap \Gamma_G(z) \neq \emptyset\} \tag{4}$$

On appelle *augmentation* du sommet z l'union d'une clique augmentante C pour z et d'une constraction de cliques de $D_z(C)$.

Un exemple de clique augmentante $C1$ pour le sommet z est donné dans la figure 1, avec $D_z(C1) = \{C2\}$. Par contre, la clique C6 ne peut pas être augmentante à cause de l'appartenance de l'arête $[u, v]$ à la clique $C7$ de C^i. Ce qui rend impossible toute contraction entre $C6$ et $C4$

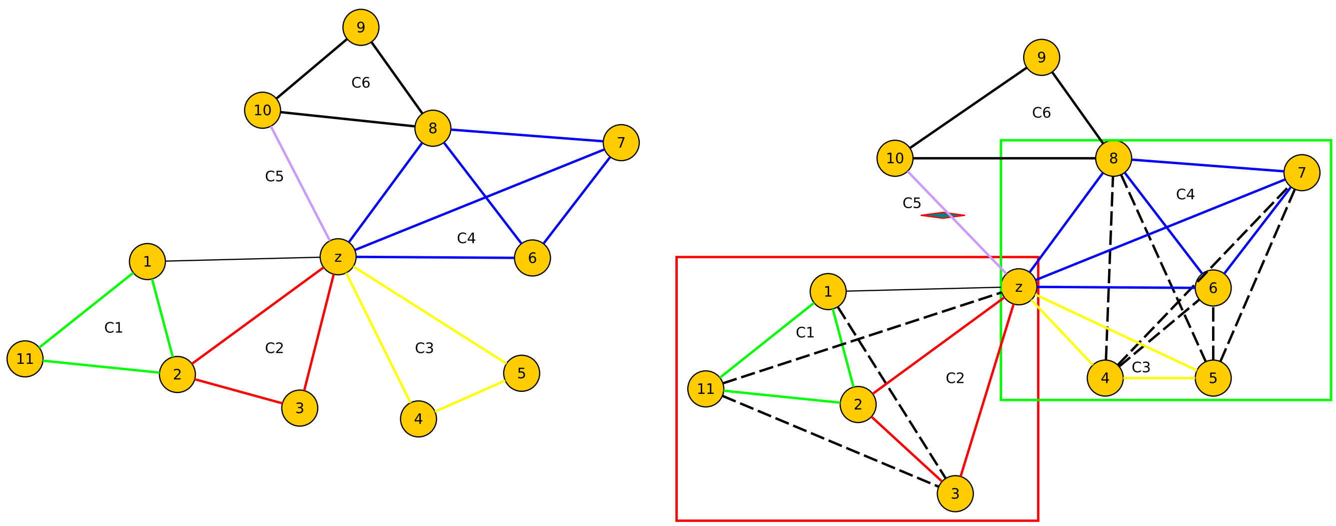

Figure 1: un exemple de compression de cliques

Définition 3 *On appelle* compression *du sommet z un triplet (π_1, π_2 et π_s) défini par :*

- *π_1 (resp. π_2) peut être chacun d'une des formes suivantes :*
  1. *l'union de z, d'un sous-ensemble C_1 (resp. C_2) de cliques de $\mathcal{C}(z)$ tel que toute paire C et C' de C_1 (resp. C_2) est contractable et d'un sous-ensemble S_1 (resp. S_2) de sommets $v \in S(z)$ n'appartenant à aucune clique de C_1 (resp. C_2) tel que*

     $$\forall v \in S_1, \ \forall x \in C_1, \nexists C' \in \mathcal{C} \ t.q. \ card(C') > 2 \ \ et \ \ \{v, x\} \subset C'$$

     (ce qui fait que $\{v,x\}$ peut etre une clique de $\mathcal{C}^i$).
  2. *une augmentation du sommet z*
- *π_1 et π_2 ne peuvent pas être simultanément réduits à $\{z\}$ et $\pi_1 \cap \pi_2 = \{z\}$,*

- $\pi_S = \Gamma_G(z) -$ $((\pi_1 \cap \Gamma_G(z)) \cup (\pi_2 \cap \Gamma_G(z)))$ *tel que l'ensemble des arêtes* $\{[z,v] \in E_C^i :\ v \in \pi_S\}$ *n'est pas déconnectant.*

- *le triplet* $\pi_1 \cap \Gamma_G(z)$, $\pi_2 \cap \Gamma_G(z)$, $\pi_S \cap \Gamma_G(z)$ *est une 3-partition de* $\Gamma_G(z)$

Il existe toujours une telle compression, ne serait-ce que $\pi_1 = \{z\} \cup C_i \in C(z)$, $\pi_2 = \emptyset$, $\pi_s = \gamma_G(z) - (\gamma_G(z) \cup C_i)$ si $\mathcal{C}(z)$ n'est pas vide. Sinon, $\pi_1 = \{z\} \cup \{v \in \gamma_G(z)\}$, $\pi_2 = \emptyset$, $\pi_s = \gamma_G(z) - \{v\}$ est aussi une compression. Un exemple de compression est aussi donné dans la figure 1. Le coût $c(T)$ d'une compression π_1, π_2, π_S est défini par :

$$c(T) = |\{\{u,v\} \in \pi_1 :\ [u,v] \notin E_C^i\}| + |\{\{u,v\} \in \pi_2 :\ [u,v] \notin E_C^i\}| + \ |\pi_S|$$

Dans l'exemple de la figure 1(a), autour d'un sommet z, l'ensemble $C(z)$ contient les cliques $C2$, $C3$,$C4$ et $C5$. Les cliques $C5$ et $C4$ ne sont pas contractables, à cause de l'existence de $C6$ dans $\mathcal{C}_i$. La clique $C1$ est voisine de z et $D(C1) = \{C2\}$. L'exemple de compression qui est donné dans la figure 1(b) est $\pi_1 = C1 \cup C2$ (une augmentation), $\pi_2 = C3 \cup C4$ (ces deux cliques étant contractables), et $\pi_s = \{x\}$. Le coût de cette compression est 10, 10 étant le nombre d'arêtes en pointillé plus l'arête supprimée $[x, z]$.

Soit $Cout(z)$ le coût minimum d'une compression de z. Le but est de modifier G_C afin que z puisse être couvert par une ou deux cliques issues de π_1 et π_2. Pour cela, le coût de cette modification $c(T)$ tient compte des arêtes à ajouter (liées à π_1 et π_2) et à supprimer (liées à π_s). Ainsi, **appliquer une compression** $T = \pi_1, \pi_2, \pi_s$ consiste à ajouter dans E_C^i les arêtes définies par les ensembles de paires $\{\{u,v\} \in \pi_1 :\ [u,v] \notin E_C^i\}$ (qui seront couvertes par la clique π_1) et $\{\{u,v\} \in \pi_2 :\ [u,v] \notin E_C^i\}$ (qui seront couvertes par la clique π_2) et à supprimer les arêtes $\{[z,v] \in E_C^i :\ v \in \pi_S\}$.
Des lors, le sommet z appartient aux deux cliques π_1 et π_2. On procède alors aux mises à jour suivantes pour obtenir $\mathcal{C}^{i+1}$ et E_C^{i+1} :

- supprimer toutes les cliques $\mathcal{C}_z$ couvertes par π_1 dans $\mathcal{C}^i$.

- supprimer toutes les cliques $\mathcal{C}_z$ couvertes par π_2 dans $\mathcal{C}^i$.

- supprimer toutes les cliques de cardinalité 2 couvertes par π_1 et π_2 dans $\mathcal{C}^i$.

- ajouter π_1 et π_2 dans $\mathcal{C}^i$, supprimer de E_C^{i+1} toutes les arêtes $\{[z,v] \in E_C^i :\ v \in \pi_S\}$.

- Affecter $Cliq(z)$ à 1 (si π_1 ou π_2 est vide) ou 2 (sinon).

Cette procédure a les propriétés suivantes :

Propriété 1 *Considérons une application d'une compression, Soit* $\mathcal{C}^{i+1}$ *l'ensemble obtenu à partir de* $\mathcal{C}^i$ *après mise à jour selon cette application.*

- *Tout sommet de* G_C *couvert par une ou deux cliques dans* $\mathcal{C}^i$ *le reste dans* $\mathcal{C}^{i+1}$.

- *Toute arête couverte par une et une seule clique dans* $\mathcal{C}^i$ *et qui n'est pas supprimée le reste dans* $\mathcal{C}^{i+1}$.

- *Le sommet* z *est couvert par une ou deux cliques dans* $\mathcal{C}^{i+1}$ *(le nombre de sommets ainsi couverts augmente de 1 par rapport à celui dans* $\mathcal{C}^i$*).*

Ainsi, pour chaque sommet z_i pris dans l'ordre O_z, on considère une compression de coût minimum c_m^i et on l'applique. La propriété ci-dessus garantit qu'à l afin du processus, on obtientun graphe de corrélation $G_C^t = (V, E_C^t)$ dont l'ensemble $\mathcal{C}$ modifié est une couverture de corrélation. La distance-line vérifie

$$DL(G_C^0, G_C^t) \leq \sum_{1 \leq i \leq t} c_m^i$$

Notons que lors d'une étape $j > 1$, le sommet z_j et son voisinage se retrouve être couvert par une ou deux cliques suite au traitement des $j-1$ sommets précédents, aucune compression ne lui est appliquée (on considère la compression identité) et donc $c_m^i = 0$.