Algorithmes pour la physique statistique

1. Modèle d'Ising
2. Modèle des dimères
3. Le cas planaire : calcul par déterminant
4. Modèle "hard-core" et coloriage
5. Algorithme d'énumération de coloriages

Algorithmes pour l'économie

Modèle d'Ising

Graphe G, interaction $J(e)$ sur chaque arête e, configuration : spin ± 1 à chaque sommet distribué selon Gibbs:

$$\Pr\{\sigma\} = exp(-\beta H(\sigma))/Z,$$

$\beta = 1/kT$, Z fonction de partition,

$$H(\sigma) = \sum_{x\tilde{y}} J(\{x,y\})\sigma_x\sigma_y.$$

Si $J < 0$ est constant : configuration la plus probable = MAXCUT

Modèle des dimères

Modélise l'adsorption de molécules di-atomiques sur une surface cristalline.

Graphe G, interaction $J(e)$ sur e, configuration : couplage parfait des sommets de G, distribué selon Gibbs: Si G 2-colorié noir-blanc, couplage parfait : bijection $\sigma : N \to B$

$$\Pr\{\sigma\} = exp(-\beta H(\sigma))/Z,$$

$\beta = 1/kT$, Z fonction de partition,

$$H(\sigma) = \sum_{x \in N} J(\{x, \sigma(x)\}).$$

$$H(\sigma) = \prod_{e \in \sigma} \lambda_e.$$

Le cas planaire : la grille

G : grille $2n \times 2n$, $J(e) = J$ constant : calculer le nombre de (ou poids des) couplages parfaits de G : problème du **permanent.**
$Z = \sum_\sigma \prod_{e\in\sigma} \lambda_e$.
Soit $B = \{1_B, 2_B, \ldots, N_B\}$ et $R = \{1_R, 2_R, \ldots, N_R\}$ avec $|B| = |R| = 2n^2$.
Soit

$$a_{ij} = \begin{cases} \lambda_e & \text{si } e = \{i_B, j_R\} \\ 0 & \text{sinon} \end{cases}$$

Soit

$$w_{ij} = \begin{cases} a_{ij} & \text{si } \{i_B, j_R\} \text{ horizontale} \\ a_{ij} & \text{si } \{i_B, j_R\} \text{ verticale en position paire} \\ -a_{ij} & \text{si } \{i_B, j_R\} \text{ verticale en position impaire} \end{cases}$$

Alors $Z = \mid det(w_{ij}) \mid$.

Analyse

Cas simplifié: $\lambda_e = 1$.

$$det(w_ij) = \sum_{\sigma \in S_n} \epsilon(\sigma) w_{1\sigma(1)} \dots w_{N\sigma(N)}.$$

Termes non nuls: couplages de la grille.

Chaque couplage parfait contribue ± 1.

il faut montrer que

$$|\sum_{\sigma \in S_n} \epsilon(\sigma) w_{1\sigma(1)} \dots w_{N\sigma(N)}| = \sum_{\sigma \in S_n} |\epsilon(\sigma) w_{1\sigma(1)} \dots w_{N\sigma(N)}|,$$

autrement dit, que tous les termes ont même signe.

Modèle "hard-core" de coloriage de graphes

Graphe G de degré maximum Δ, k couleurs, coloriage $\sigma : V \to \{1, \dots k\}$, probabilité proportionnelle à

$$\prod_{e=\{i,j\}} f(\sigma(i), \sigma(j)).$$

Calculer $Z = \sum_\sigma \prod_{e=\{i,j\}} f(\sigma(i), \sigma(j))$.

Cas hard-core: $f(c, c') = 1$ si $c \neq c'$, et $= 0$ sinon.

Distribution uniforme sur les k-coloriage valides

Générer un k-coloriage au hasard.

Algorithme

Hypothèse: $k \geq 2\Delta + 1$.

Initialisation : $\sigma_0 \leftarrow$ un coloriage arbitraire.

Répéter $O(n \log n)$ fois:

- Soit $v \in V$ choisi au hasard uniforme
- Soit $c \in \{1, \dots, k$ choisie au hasard uniforme
- Si colorier v avec c est valide, le faire

Sortie : le coloriage courant

Compter les k-coloriages

Soient $e_1, \dots e_m\}$ les arêtes de G, Soit G_i le graphe contenant les arêtes $\{e_1, \dots e_i\}$. Soit $Z(G_i)$ le nombre de k coloriages de G_i. Alors:

$$Z(G) = \prod_i \frac{Z(G_i)}{Z(G_{i-1})} \times Z(G_0).$$

Pour estimer $Z(G)$:

$Z(G_0) = k^n$.

Pour estimer $Z(G_i)/Z(G_{i-1})$: générer des coloriages aléatoires de G_{i-1}, estimer la probabilité qu'ils soient également coloriages valides de G_i.

Analyse:

$$\frac{1}{3} \leq \frac{Z(G_i)}{Z(G_{i-1})} \leq 1$$

d'où estimation fiable avec peu d'échantillons

Algorithmes pour l'économie

Analyse d'enchères.

Interaction entre un acheteur et un groupe de vendeurs d'un objet particulier, ou entre un vendeur et un groupe de clients potentiels. Stratégies à concevoir: quel montant le client doit il proposer à l'enchère pour obtenir l'objet ou pour maximiser son gain? A quel prix le vendeur doit il mettre en vente l'objet?

1. Enchères de Vickrey
2. mécanisme de VCG appliqué aux plus courts chemins ou à l'arbre couvrnt minimum
3. Frugalité: coût de l'honnêteté.

Enchères

Quelle entreprise choisir pour construire le nouveau bâtiment?

Coûts : n entreprises, l'entreprise i peut faire le travail à un coût c_i (minimum à payer pour que l'entreprise fasse un profit positif). c_i n'est connu que de l'entreprise (valeur privée).

Soit p_i le montant qu'on paye à l'entreprise i. L' **utilité** du travail pour i est:

$$u_i = \begin{cases} 0 & \text{si } i \text{ ne gagne pas} \\ p_i - c_i & \text{si } i \text{ gagne l'enchère} \end{cases}$$

But de i: maximiser u_i.
Pas de collusions.

Enchères cachées

1. Chaque joueur (entreprise) donne une enveloppe scellée au vendeur.
2. Le vendeur ouvre toutes les enveloppes simultanément et décide quel joueur gagne et à quel prix.

Enchères les plus simples: le gagnant correspond à l'enchère minimale, et c'est là le prix qu'il reçoit. Problème: les joueurs ne savent pas combien ils devraient proposer – probablement plus que leur coût c_i.

Enchères de Vickrey

Le gagnant est le joueur dont l'enchère est minimale, mais le vendeur le paie un montant égal à la **deuxième meilleure enchère**

L'honnêteté est alors la meilleure stratégie pour chaque joueur.

Théorème: Indépendamment de ce que choisissent les autres joueurs, la meilleure stratégie pour le joueur i consiste à proposer son travail au prix c_i.

Preuve. Soit i fixé. Soit T la meilleure enchère parmi les joueurs $j \neq i$: $T = \min\{b_j, 1 \leq j \leq n, j \neq i\}$.

Cas 1: si $c_i < T$, alors, si le joueur i dit c_i, il gagne, il sera payé T, et son utilité sera $T - c_i$. Si le joueur dit $b_i \neq c_i$, il pourrait gagner ou perdre; s'il gagne il sera toujours payé T et son utilité sera toujours $T - c_i$. S'il perd son utilité sera $0 < T - c_i$, donc il n'a pas intérêt à mentir dans ce cas.

Cas 2: si $c_i > T$, alors, si le joueur i dit c_i, il perd et $u_i = 0$. S'il dit $b_i > c_i$, idem. S'il dit $b_i < c_i$, il gagne mais est payé T et son utilité est $T - c_i < 0$, donc il n'y a oas intérêt.

Remarques

- Choix du joueur réalisant l'"optimum social": c_i minimum, donc entreprise utilisant le minimum de ressources naturelles.
- Il existe une valeur seuil T indépendante de i telle que le joueur i gagne si et seulement si son enchère est meilleure que T
- Le paiement fait au gagnant i égale son coût déclaré b_i plus sa valeur marginale. Valeur marginale: différence entre le coût si i ne participait pas et le coût déclaré de i.

Plus court chemin

Graphe G, deux noeuds s et t, chaque arête appartient à un joueur différent. Le poids (privé) d'une arête est le coût c_e d'acheminement d'un message de long de l'arête e. Comment décider quel chemin utiliser pour envoyer un message de s à t, et combien payer à chaque arête du chemin?

Minimum: Longueur du plus court chemin dans G valué par $(c_e)_e$.

Problème: les c_e sont des valeurs privées...

Mécanisme de Vickrey-Clarke-Grove (VCG)

1. Chaque joueur (arête) donne une enveloppe scellée contenant la valeur b_e à laquelle il met son arête e à prix.
2. On ouvre simultanément toutes les enveloppes, on calcule le plus court chemin dans G avec les poids $(b_e)_e$.
3. On paie à chaque joueur e se trouvant sur le plus court chemin un montant égal à

$$p_e = b_e + \mathrm{PCC}(G - e) - \mathrm{PCC}(G).$$

Honnêteté

La meilleure stratégie pour e consiste à annoncer $b_e = c_e$.

Problèmes

Exemple : Multigraphe G à deux sommets (s,t) et deux arêtes de coûts 1 et 100: le gagnant déclare 1 mais est payé 100... le coût est très élevé par rapport à ce qui serait possible si les valeurs des arêtes n'étaient pas privées!

Malheureusement : Pour le problème du plus court chemin, ces surcoûts sont impossibles à éviter si on veut garantir l'honnêteté des joueurs.

Arbre couvrant minimum

1. Chaque joueur (arête) donne une enveloppe scellée contenant la valeur b_e à laquelle il met son arête e à prix.
2. On ouvre simultanément toutes les enveloppes, on calcule le MST dans G avec les poids $(b_e)_e$.
3. On paie à chaque joueur e se trouvant sur le MST un montant égal à

$$p_e = b_e + \mathrm{MST}(G-e) - \mathrm{MST}(G).$$

Théorème : Le coût total est au plus celui de $T' = \mathrm{MST}(G \setminus T^*)$, où T^* est le MST de G avec les coûts déclarés.

Analyse

Lemme: Si k arêtes $\{e_1, \ldots, e_k\}$ de T^* reçoivent paiement $\geq p$, alors il existe k arêtes de T' de coût $\geq p$.

Preuve: Le graphe $G - \{e_1 \ldots, e_k\}$ a $k+1$ composantes connexes. Soit e une arête de G entre deux des composantes. Montrons que son coût est $\geq p$.

$T^* \cup e$ contient un cycle, on coupe l'arête e_j de T^* qui relie les deux composantes, et soit $U = T^* + e - e_j$. Le paiement fait à l'arête e_j est:

$$\begin{aligned} p \leq c_{e_j} + \mathrm{MST}(G-e_j) - \mathrm{MST}(G) &\leq c_{e_j} + c(U) - c(T^*) \\ &\leq c(e_j) + c(e) - c(e_j) \\ &\leq c(e) \end{aligned}$$

D'où le lemme.

Analyse (suite).

Soient $p_{e_1} \geq \ldots \geq p_{e_{n-1}}$ les prix payés aux arêtes de T^*. Soient $c_{f_1} \geq \cdots \geq c_{f_{n-1}}$ les coûts des arêtes de T'. Par le lemme, $c_{f_i} \geq p_{e_i}$ pour tout i, donc le prix total payé est au plus $c(T')$.