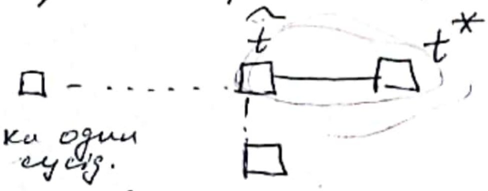
Для довільного напівкільця $\langle R, \oplus, \otimes \rangle$, ациклічного орієнтованого графу $\langle T, \tau \rangle$, відомої скінченної множини K і функції $g: \tau \times K^2 \to R$ навести алгоритм (ефективний) пошуку значення:

$$G = \bigoplus_{k \in K^T} \bigotimes_{tt' \in \tau} g_{tt'}(k_t, k_{t'})$$



Знайдеться об'єкт в якого тільки один сусід.

$\exists! \hat{t} \in T : \hat{t} \in N(t^*)$ - тобто $\hat{t}, t^*$ - сусіди.

$$G = \bigoplus_{k \in K^T} \bigotimes_{tt' \in \tau} g_{tt'}(k_t, k_{t'}) = \Big\{ \text{розділимо } \hat{t} \text{ і } t^* \Big\} =$$

$$= \bigoplus_{\bar{k} \in K^{T \setminus t^*}} \bigoplus_{k^* \in K} \bigotimes_{tt' \in \tau \setminus \{\hat{t}t^*, t^*\hat{t}\}} g_{tt'}(k_t, k_{t'}) \otimes g_{\hat{t}t^*}(k_{\hat{t}}, k_{t^*}) =$$

$$= \left\{ \begin{array}{l} \text{введемо позначення: } T_2 = T \setminus t^* \\ \tau_2 = \tau \setminus \{\hat{t}t^*, t^*\hat{t}\} \end{array} \right\} =$$

$$= \bigoplus_{\bar{k} \in K^{T_2}} \bigoplus_{k^* \in K} \bigotimes_{tt' \in \tau_2} g_{tt'}(k_t, k_{t'}) \oplus g_{\hat{t}t^*}(k_{\hat{t}}, k_{t^*}) =$$

$$= \{ \text{викор. дистрибутивність} \} =$$

$$= \bigoplus_{\bar{k} \in K^{T_2}} \bigotimes_{tt' \in \tau_2} g_{tt'}(k_t, k_{t'}) \bigoplus_{k^* \in K} g_{\hat{t}t^*}(k_{\hat{t}}, k_{t^*})$$

Для $\forall k_{\hat{t}} \in K$:

$$g_{\hat{t}\hat{t}}(k_{\hat{t}}, k_{\hat{t}}) \circledast = \bigoplus_{k^* \in K} g_{\hat{t}t^*}(k_{\hat{t}}, k_{t^*})$$

Виключаємо об'єкти один за одним.

$$\bigoplus_{k^* \in K} g_{tt^*}(k_t, k_{t^*})$$

Чому знайдеться такий об'єкт у якого тільки один сусід?

Тобто, чи залишиться дерево деревом, якщо прибрати з нього лист?

Дерево - зв'язний ациклічний граф.

Лист поєднаний з деревом одним ребром.

Коли ми вилучаємо лист ми вилучаємо це ребро.

=> отримали граф.
=> Перевіримо чи є цей граф деревом.
Зв'язність не порушено, бо не існує жодного маршруту який проходить через лист, що з'єднує дві інші вершини.
Очевидно, що не можна створити цикл коли вилучаємо ребра ▲

2. Для ациклічного орієнтованого графу (T, τ), відомої скінченної множини K і функції $g: \tau \times K^2 \to [0,1]$ навести ефективний алгоритм пошуку значення.

$$G = \sum_{k \in K^T} \prod_{tt' \in \tau} g_{tt'}(k_t, k_{t'}) \qquad \exists! \hat{t} \in T : \hat{t} \in N(t^*).$$

$$G = \sum_{k \in K^T} \prod_{tt' \in \tau} g_{tt'}(k_t, k_{t'}) = \left\{ \text{розділимо } \hat{t} \text{ і } t^* \right\} =$$

$$= \sum_{\bar{k} \in K^{T \setminus k^*}} \sum_{k^* \in K} \prod_{tt' \in \tau \setminus \{\hat{t}t^*, t^*\hat{t}\}} g_{tt'}(k_t, k_{t'}) \cdot g_{\hat{t}t^*}(k_{\hat{t}}, k_{t^*})$$

$$= \left\{ \text{введемо позначення} : T_2 = T \setminus t^*,\ \tau_2 = \tau \setminus \{\hat{t}t^*, t^*\hat{t}\} \right\} =$$

$$= \sum_{\bar{k} \in K^{T_2}} \sum_{k^* \in K} \prod_{tt' \in \tau_2} g_{tt'}(k_t, k_{t'}) \cdot g_{\hat{t}t^*}(k_{\hat{t}}, k_{t^*}) =$$

$$= \left\{ \text{викор. дистрибутивність} \right\} =$$

$$= \sum_{\bar{k} \in K^{T_2}} \prod_{tt' \in \tau_2} g_{tt'}(k_t, k_{t'}) \sum_{k^* \in K} g_{\hat{t}t^*}(k_{\hat{t}}, k_{t^*})$$

$$\forall\ k_{\hat{t}} \in K$$

$$g_{\hat{t}\hat{t}}(k_{\hat{t}}, k_{\hat{t}}) = g_{\hat{t}t^*}(k_{\hat{t}}, k_{t^*})$$

Виключаємо об'єкти один за одним.

$$\sum_{k^* \in K} g_{tt^*}(k_t, k_{t^*})$$

Постановка задачі та стратегія розпізнавання текстів на бінарних зображеннях з шумом Бернуллі з відомим параметром $0 \le p \le 1$ за штрафної функції $\omega(k,d) = \mathbb{1}(k \ne d)$. Шум накладається незалежно на кожен піксель за допомогою виключного „АБО", апріорна ймовірність кожного символу $p(k)$ відома, еталонні зображення $q: K \to \{0,1\}^{h \times *}$ символів відомі, ширини різних символів можуть відрізнятися, висоти h однакові.

$X = \{0,1\}^{n \times m}$ – мн-на вхідних зображень.

$K = \{T \in A^* \mid \tau(T) \in \{0,1\}^{n \times m}\}$ – мн-на слів, що $\tau(T)$ – є зображенням цього слова.

$\tau: A^* \to \bigcup\limits_{a \in N_0} \{0,1\}^{n \times a}$ – переводить текст у зображення.

A – алфавіт з ␣ – пропуск

A^* – мн-на всіх можливих текстів, що містять тільки символи з A.

Ризик:

$$\sum_{x \in X} \sum_{k \in K} p(x,k) \cdot \omega(k, q(x)) \to \min_{q: X \to K}$$

штрафна ф-я: $\omega(k,d) = \mathbb{1}(k \ne d)$

$$= \sum_{x \in X} \sum_{k \ne q(x)} p(x,k) \to \min_{q: X \to K}.$$

Застосуємо індикатор. $\Rightarrow \sum\limits_{x \in X} p(x, q(x)) \to \max\limits_{q: X \to K}.$

$x = \tau(k) \oplus N$, N – шум Бернуллі

Потрібно виділити шум $\Rightarrow x \oplus \tau(k) = (\tau(k) \oplus N) \oplus \tau(k) =$
$= \tau(k) \oplus \tau(k) \oplus N = N \Rightarrow x \oplus \tau(k) = N.$

$$p_{\xi}(x \oplus \tau(k)) = \prod_{i=1,n} \prod_{j=1,m} p^{x_{ij} \oplus \tau(k)_{ij}} \cdot (1-p)^{1 \oplus x_{ij} \oplus \tau(k)_{ij}}$$

$$p(k) = p(k_0) \cdot \prod_{i=1}^{|k|-1} p(k_i / k_{i-1}) \cdot p(k_{|k|} = ␣ / k_{|k|-1}) -$$

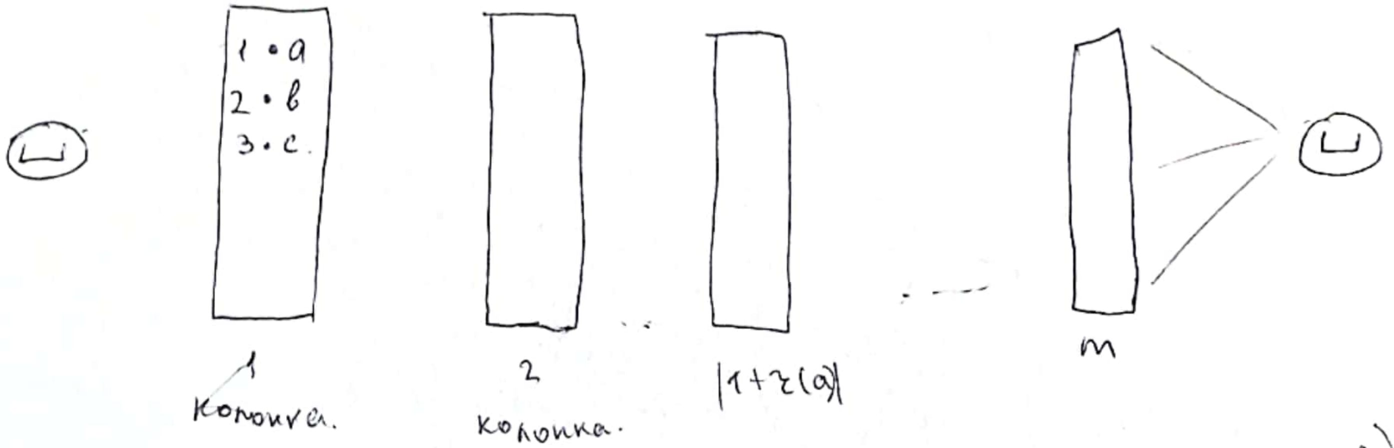
– апріорна ймов-ть k; w – ширина вхідного символу.

Є вхідний рядок розміром $n \times m$ (як зображення)

Створюємо масив f розміром $m \times |A|$.

Заповнюємо всі значення $-inf$.

$f[i][j]$ – штраф за те, що на i-ій колонці починається j-та буква; всіх букв $27 = |A|$



$$f[1,S] = \left(\sum_{i=1}^{n} \sum_{j=1}^{|r(e_s)|} \left(x_{ij} \oplus r(e_s)_{ij}\right) \cdot \ln(p) + \ln(1-p)\left(1 \oplus x_{ij} \oplus r(e_s)_{ij}\right)\right)$$

$$+ \ln\left(p(e_s)\right), \ \forall s \in (1, .., 27),$$ S – множина індексів букв. e_s – буква

$$f[t,S] = \max\left(f_{t-|r(e)|} + \left(\sum_{i=1}^{n} \sum_{j=t}^{t+|r(e)|} \dots \right) + \ln\left(p(e, e_t)\right)\right) \quad \forall e \in A.$$

$$e_1 = \underset{s}{\operatorname{argmax}} \left(f[m,S] \right).$$

$$e_2 = \underset{s}{\operatorname{argmax}} \left(f[m - |r(e_1)|; s] + \ln\left(p(e_1, s)\right)\right).$$

$\vdots$

Показати, що для будь-яких двох супермодулярних ф-ій $g_1 : K^2 \to \mathbb{R}$ і $g_2 : K^2 \to \mathbb{R}$ функція:

$g(k, k') = \max_{c \in K} \{ g_1(k, c) + g_2(c, k') \}$. Також супермодулярна

$g_1(k, c)$ - супермодулярна $\forall\ k_1, k_2, c_1, c_2$.

$g_1(k_1, c_1) + g_1(k_2, c_2) \le g_1(k_1 \vee k_2, c_1 \vee c_2) + g_1(k_1 \wedge k_2, c_1 \wedge c_2)$ *

$g_2(c, k')$ - супермодулярна $\forall\ c_1, c_2, k_1', k_2'$.

$g_2(c_1, k_1') + g_2(c_2, k_2') \le g_2(c_1 \vee c_2, k_1' \vee k_2') + g_2(c_1 \wedge c_2, k_1' \wedge k_2')$ **

* + **

$\Rightarrow g_1(k_1, c_1) + g_2(c_1, k_1') + g_1(k_2, c_2) + g_2(c_2, k_2') \le$

$g_1(k_1 \vee k_2, c_1 \vee c_2) + g_2(c_1 \vee c_2, k_1' \vee k_2') + g_1(k_1 \wedge k_2, c_1 \wedge c_2) + g_2(c_1 \wedge c_2, k_1' \wedge k_2')$.

Введемо $r(x, y, z) = g_1(x, z) + g_2(z, y)$.

$\Rightarrow r(k_1, k_1', c_1) + r(k_2, k_2', c_2) \le r(k_1 \vee k_2, k_1' \vee k_2', c_1 \vee c_2) + r(k_1 \wedge k_2, k_1' \wedge k_2', c_1 \wedge c_2)$

$\forall\ k, k', c \Rightarrow r(k, k', c)$ - супермодулярна

а значить сума супермодулярних ф-ій є супермодулярною ф-єю.

Візьмемо $\max_{c_1}$, $\max_{c_2}$ від обох частин:

$$\Rightarrow \max_{c_1} (r(k_1, k_1', c_1)) + \max_{c_2} (r(k_2, k_2', c_2)) \le$$

$$\le \max_{a} (r(k_1 \vee k_2, k_1' \vee k_2', a)) +$$

$$\max_{b} (r(k_1 \wedge k_2, k_1' \wedge k_2', b))$$

а обозначим через $c_1 \vee c_2$

в 1-1 $c_1 \wedge c_2$.

if $f(x) \le f(y) \Rightarrow \max_x f(x) \le \max_y f(y)$

Показати, що для $\forall$ увігнутої $f: \mathbb{R} \to \mathbb{R}$ та скінченної множини точок міток $K \subset \mathbb{R}$ ф-я: $g(k, k') = f(k - k')$

$g(k, k')$ — є супермодулярною, коли

$\forall\, k_1, k_2, k_1', k_2'$; $\boxed{k_1 > k_2,\ k_1' < k_2'}$ - випадок.

$$g(k_1, k_1') + g(k_2, k_2') \leq g(k_1, k_2') + g(k_2, k_1')$$

перепишемо в f:

$$f(k_1 - k_1') + f(k_2 - k_2') \leq f(k_1 - k_2') + f(k_2 - k_1')$$

$$\left.\begin{aligned} a &= k_1 - k_1' \\ b &= k_2 - k_2' \\ c &= k_1 - k_2' \\ d &= k_2 - k_1' \end{aligned}\right\} \Rightarrow f(a) + f(b) \leq f(c) + f(d)\ ?$$

$**$

Випишемо вл-ть увігнутості f:

$$f(\alpha_1 a + \beta_1 b) \geq \alpha_1 f(a) + \beta_1 f(b)\ ;\ \alpha_1 \in [0, 1]$$

$$f(\alpha_2 a + \beta_2 b) \geq \alpha_2 f(a) + \beta_2 f(b)\ ;\ \alpha_2 \in [0, 1]$$

$$f(\alpha_1 a + \beta_1 b) + f(\alpha_2 a + \beta_2 b) \geq (\alpha_1 + \alpha_2) f(a) + (\beta_1 + \beta_2) f(b)$$

$*$

$$\begin{cases} \alpha_1 + \beta_1 = 1. \\ \alpha_1 + \beta_1 b = c \end{cases}$$ — хочемо показати.

$$\begin{cases} \beta_1 = 1 - \alpha_1. \\ \alpha_1 a + b - \alpha_1 b = c \end{cases} \Rightarrow \alpha_1 (a - b) = c - b \Rightarrow \alpha_1 = \frac{c - b}{a - b},\ \beta_1 = \frac{a - c}{a - b}$$

$$\begin{cases} \alpha_2 + \beta_2 = 1 \\ \alpha_2 a + \beta_2 b = d \end{cases} \Rightarrow \alpha_2 = \frac{d - b}{a - b},\ \beta_2 = \frac{a - d}{a - b}$$

Для того, щоб $* = **$, потрібно показати, що:

$\alpha_1 + \alpha_2 = 1.$

$$\Rightarrow \frac{c - b}{a - b} + \frac{d - b}{a - b} = \frac{c - b + d - b}{a - b} =$$

$$= \frac{a + b - b - b}{a - b} = 1.$$

Аналогічно для $\beta_1 + \beta_2$.

$$\Rightarrow f(c) + f(d) \geq$$

$$\geq f(a) + f(b)$$

Для $k_1 > k_2$, $k_1' < k_2'$

$a + b = c + d.$

Вивести алгоритм дифузії для покращення оцінки зверху ймовірності найбільш ймовірного стану випадкового поля маркова.

$$\min_{\varphi: T \to K} \left\{ \sum_{t \in T} \max_{k \in K} \left(q_t(k) - \sum_{t' \in N(t)} \varphi_{tt'}(k) \right) + \sum_{tt' \in T} \max_{k, k' \in K} \left(g_{tt'}(k,k') + \varphi_{tt'}(k) + \varphi_{t't}(k') \right) \right\}$$

Складність задачі полягає в тому, що вона дискретна, тобто нам необхідно для кожного об'єкта вибрати по одній мітці (строго один з варіантів). Ми маємо позбутись цієї дискретності. Постараємся тепер для кожного об'єкту обрати якусь суміш із міток.

Нехай є: $\alpha_t(k_t)$, $t \in T$, $k_t \in K$ - ваги для кожної мітки.

$\beta_{tt'}(k_t, k_{t'})$, $tt' \in T$, $k_t \in T$, $k_t \in K$, $k_{t'} \in K$ - ваги дуг.

$t \in T$ - множина об'єктів, $tt' \in T$ - множина сусідів, $N(t) = \{t' \mid tt' \in T\}$.

$q_t(k)$ - унарний штраф, $g_{tt'}(k,k')$ - бінарний штраф.

$\varphi_{tt'}(k)$, $\varphi_{t't}(k')$ - репараметризація. $\boxed{\text{Зараз } F(k) = \sum_{t \in T} q_t(k) + \sum_{tt' \in T} g_{tt'}(k_t k_{t'})}$

Запишемо задачу лінійного програмування:

$$\sum_{t \in T} \sum_{k \in K} \alpha_t(k) q_t(k) + \sum_{tt' \in T} \sum_{k_1 \in K} \sum_{k_2 \in K} \beta_{tt'}(k_1, k_2) \cdot g_{tt'}(k_1, k_2) \xrightarrow{\alpha, \beta} \max.$$

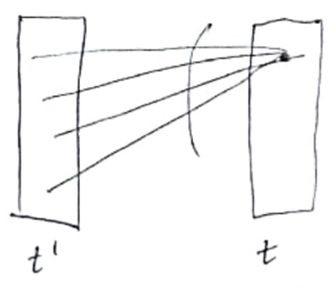
$\sum_{k \in K} \alpha_t(k) = 1$, $t \in T$ - Ваги сумуються в 1. | Θ_t

$\sum_{k_1 \in K} \sum_{k_2 \in K} \beta_{tt'}(k_1, k_2) = 1$, $tt' \in T$ - Ваги дуг теж сумуються в 1. | $\gamma_{tt'}$

$\alpha_t(k) - \sum_{k_2 \in K} \beta_{tt'}(k_1, k_2) = 0$, $t = T$, $t' \in N(t)$, $k \in K$ | $\varphi_{tt'}(k)$.

$\alpha_t(k) \geq 0$, $t \in T$, $k \in K$

$\beta_{tt'}(k_1, k_2) \geq 0$, $tt' \in T$, $k_1 \in K$, $k_2 \in K$



Сума ваг всіх цих дуг = вазі мітки в об'єкті t.

Переходимо до двоїстої задачі:

$$\sum_{t \in T} \Theta_t + \sum_{tt' \in T} \gamma_{tt'} \xrightarrow{\Theta, \gamma, \varphi} \min.$$

$\alpha_t(k)$: $\Theta_t + \sum_{t' \in N(t)} \varphi_{tt'}(k) \geq q_t(k)$

$\beta_{tt'}(k_1, k_2)$: $\gamma_{tt'} - \varphi_{tt'}(k_1) - \varphi_{t't}(k_2) \geq g_{tt'}(k_1, k_2)$

$tt' \in T$, $k_1 \in K$, $k_2 \in K$.

Звідси отримуємо, що:
Ми хочемо зробити Θ якомога меншим, але більшим за
$q_t(k) - \sum_{t'\in N(t)} \varphi_{tt'}(k), \; t \in T, k \in K.$

$$\Rightarrow \Theta_t = \max_{k\in K}\Big(q_t(k) - \sum_{t'\in N(t)} \varphi_{tt'}(k)\Big), \; t\in T, k \in K.$$

Аналогічно, для $g_{tt'} \Rightarrow g_{tt'} = \max_{\substack{k_1\in K\\ k_2\in K}} \big(g_{tt'}(k_1,k_2) + \varphi_{tt'}(k_1) + \varphi_{t't}(k')\big)$

$\Rightarrow$ Підставимо в $F(k)$

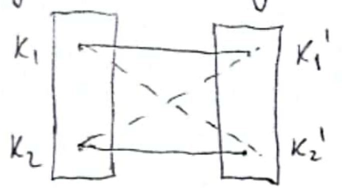
$$\Rightarrow \sum_{t\in T} \max_{k\in K}\Big(q_t(k) - \sum_{t'\in N(t)} \varphi_{tt'}(k)\Big) + \sum_{tt'\in\tau} \max_{\substack{k_1\in K\\ k_2\in K}} \big(g_{tt'}(k_1,k_2) + \varphi_{tt'}(k_1) + \varphi_{t't}(k_2)\big) \xrightarrow{\varphi} \min$$

Супермодулярність

Ми маємо справу з мітками з K. Нехай K - повністю впорядковане
$\Rightarrow \forall k_1, k_2 \in K$: або $k_1 \geq k_2$ або $k_2 \geq k_1$

Визначення супермодулярності:
$g(k,k')$ — супермодулярна, якщо:
$\forall k_1, k_2, k_1', k_2' \in K'$

$$g(k_1,k_1') + g(k_2,k_2') \leq g(k_1 \vee k_2, k_1' \vee k_2') + g(k_1 \wedge k_2, k_1' \wedge k_2')$$



"Сума якостей двох дуг, що перетинаються
$\leq$ ніж сума якостей двох паралельних дуг"

Чи впливає репараметризація на супермодулярність?

Ні.

Нехай ми застосували деяку репараметризацію і отримали нову ф-ю.

$$g'(k,k') = g(k,k') - \varphi_{tt'}(k) - \varphi_{t't}(k')$$

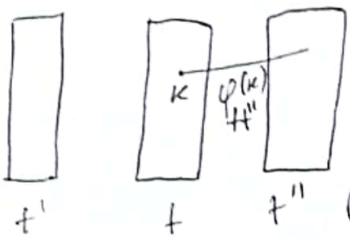
Для супермодулярності має виконуватись наступне

$$g(k_1,k_1') - \cancel{\varphi_{tt'}(k_1)} - \cancel{\varphi_{t't}(k_1')} + g(k_2,k_2') - \cancel{\varphi_{tt'}(k_2)} - \varphi_{t't}(k_2') \leq$$
$$g(k_1,k_2') - \cancel{\varphi_{tt'}(k_1)} - \cancel{\varphi_{t't}(k_2')} + g(k_2,k_1') - \cancel{\varphi_{tt'}(k_2)} - \varphi_{t't}(k_1')$$

Алгоритм дифузії

Важлива властивість алгоритму на кожній ітерації?

Для всіх $\varphi_{tt'}(k)$, $t' \in N(t)$, $k \in K$.



Ми намагаємось знайти мінімум G.

Дифузія намагається знайти таку репараметризацію

(Підібрати таке значення φ, щоб якість максимальних дуг які виходять з t в tt'' і t' було однаково)

тобто дифузія намагається:

$$\max_{k' \in K}\left(g_{tt'}(k,k') - \varphi_{tt'}(k) - \varphi_{t't}(k')\right) =$$

$$= \max_{k'' \in K}\left(g_{tt''}(k,k'') - \varphi_{tt''}(k) - \varphi_{t''t}(k'')\right), \quad t', t'' \in N(t).$$

Ми хочемо змінити якимось чином значення φ в об'єкті t:

$$\begin{cases} q_t(k) + \sum\limits_{t' \in N(t)} \varphi_{tt'}(k) = 0. \\ \max\limits_{k' \in K}\left(g_{tt'}(k,k') - \varphi_{tt'}(k) - \varphi_{t't}(k')\right) = C = const. , \quad t' \in N(t). \end{cases}$$

Ми хочемо змінити тільки ці φ

$$k^*(t') = \arg\max_{k' \in K}\left(g_{tt'}(k,k') - \varphi_{tt'}(k) - \varphi_{t't}(k')\right).$$

додали 0

Просумуємо всі якості

$$\sum_{t' \in N(t)}\left(g_{tt'}(k,k^*(t')) - \varphi_{tt'}(k) - \varphi_{t't}(k^*(t'))\right) + q_t(k) + \sum_{t' \in N(t)} \varphi_{tt'}(k)$$

$|N(t)|$ раз просумували C.

$$= \sum_{t' \in N(t)} g_{tt'}(k,k^*(t')) - \sum_{t' \in N(t)} \varphi_{t't}(k^*(t')) + q_t(k) = |N(t)| \cdot C + 0.$$

$$\Rightarrow C = \frac{\sum\limits_{t' \in N(t)} g_{tt'}(k,k^*(t')) - \sum\limits_{t' \in N(t)} \varphi_{t't}(k^*(t')) + q_t(k)}{|N(t)|}$$

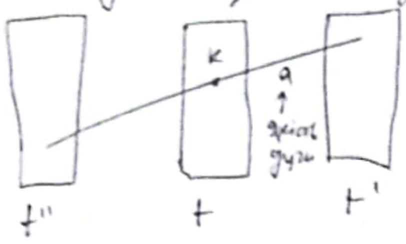
Знайшли C:

$$g_{tt'}(k,k^*(t')) - \varphi_{tt'}(k) - \varphi_{t't}(k^*(t')) = C$$

$$\Rightarrow \varphi_{tt'}(k) = g_{tt'}(k,k^*(t')) - \varphi_{t't}(k^*(t')) - C$$

Після цього всі мітки = 0, всі max дуги рівні між собою.

Розглянемо властивість, яка виконується для всіх задач в якому алгоритм дифузії припинив свою роботу!

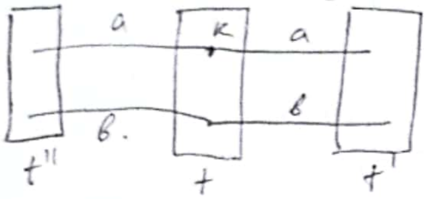


"Кожне max ребро можна продовжити на max ребро". $a = \max_{\substack{k \in K \\ k' \in K}} g_{tt'}(k,k')$

Тобто.

Якщо із t в t' серед всіх дуг a - максимальна, то із t в t'' серед всіх максимальних дуг знайдеться така дуга, що теж виходить з мітки k.

▲



Нехай a-max із $t \to t'$

$\Rightarrow$ a max із t в t'', бо max дуги із k рівні.

Але нехай $в > a$ із t в t'' $\Rightarrow$ після дифузії із t в t' теж $в$-max $\Rightarrow$ але це не так, бо ми з самого початку сказали, що a-max.

▲

Ітерації алгоритму дифузії

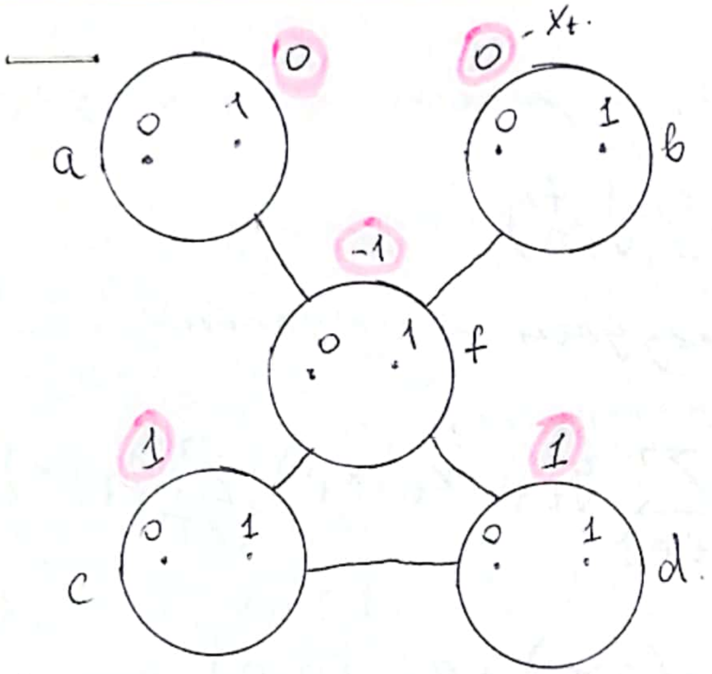
$\forall t \in T, k \in K.$

$$k^*(t') \in \arg\max_{k_{t'} \in K} \left[g_{tt'}(k_t, k_{t'}) - \varphi_{tt'}(k_t) - \varphi_{t't}(k_{t'})\right], \quad t' \in N(t)$$

$$C(t) = \sum_{t' \in N(t)} g_{tt'}(k, k^*(t')) - \sum_{t' \in N(t)} \varphi_{t't}(k^*(t')) + q_t(k_t)$$

$$\varphi_{tt'}(k) = g_{tt'}(k, k^*(t')) - \varphi_{t't}(k^*(t')) - C(t)$$

Розмітка $\forall$ сусіда $\arg\max_{k} \left(\max_{k'} \left(g_{tt'}(k,k') - \varphi_{tt'}(k) - \varphi_{t't}(k')\right)\right)$



У нас.

$g_{tt'}(k,k') = 1\!\!1\,(k=k')$

$g_t(k) = 1\!\!1\,(k = X_t)$.

предпочтение. (интерес).

$\forall t \in T, k \in K$.

$$k^*(t') \in \underset{k_{t'} \in K}{\operatorname{argmax}} \left[g_{tt'}(k_t, k_{t'}) - \varphi_{tt'}(k_t) - \varphi_{t't}(k_{t'}) \right], t' \in N(t)$$

$$C(t) = \frac{\sum\limits_{t' \in N(t)} g_{tt'}(k, k^*(t')) - \sum\limits_{t' \in N(t)} \varphi_{t't}(k^*(t')) + g_t(k_t)}{|N(t)|}$$

$$\varphi_{tt'}(k) = g_{tt'}(k, k^*(t')) - \varphi_{t't}(k^*(t')) - C(t)$$

Група людей разбилась по интересам.
Ребра – это связь „общение" между человеками.
Т. е. „~~а~~" общается с „~~b~~f", „с" с „~~f~~"

Человек хочет преднадлежать к какой-то группе.

Есть 3 группы: 0, 1, -1 – предпочтение.

Упражнение 1.

Посчитать для определенной разметки штр[аф]
этой системы.

Дано: разметка: $\overset{\quad a\ \ b\ \ c\ \ d\ \ f}{k^* = (0,0,0,0,0)}$.

(значит, что мы используем все нулевые
метки каждого объекта).

штраф (энергия): $E(\bar{k}) = \sum\limits_{tt' \in \tau} g_{tt'}(k_t, k_{t'}) + \sum\limits_{t \in T} q_t(k)$

Решение:

$$E(k^*) = g_{fa}(0,0) + g_{fb}(0,0) + g_{fc}(0,0) +$$

$$+ g_{fd}(0,0) + g_{cd}(0,0) + q_a(0) + q_b(0) + q_c(0) +$$

$$+ q_d(0) + q_f(0) = \underbrace{1 + 1 + 1 + 1 + 1}_{\Sigma g} + 1 + 1 + 0 + 0 + 0 =$$

$$= 7.$$

...нение 2

...читать оптимальную разметку.

...шение:

...нициализация: $\varphi(k)_{tt'} = 0, \ \forall t, t' \in \tau \cup \tau^{-1}, \ k \in K.$

$K = \{0, 1\}$.

$\boxed{\text{Считаем для „a"}}$

Для „a". 0 — для метки 0 в объекте „a":

Соседи „a": f.

$k_t = 0, \ k_{t'} \in \{0,1\}, \ t = a, \ t' = f.$

$\Rightarrow$ 1) $k^*(f) \in \underset{k_{t'} \in \{0,1\}}{\operatorname{argmax}} \big(g_{af}(\underset{k_t}{0}, \underset{k_{t'}}{0}) - \varphi_{af}(0) - \varphi_{fa}(\underset{k_{t'}}{0}), \ g_{af}(0,1) - \varphi_{af}(0) - \varphi_{fa}(1) \big) =$

$= \underset{k_{t'} \in \{0,1\}}{\operatorname{argmax}} \big(\underbrace{\mathbb{1}(0=0) - 0 - 0}_{1}, \ \underbrace{\mathbb{1}(0=1) - 0 - 0}_{0} \big) = \{0\}.$

2) $C(a) = \dfrac{g_{af}(0, k^*(f)) - \varphi_{fa}(k^*(f)) + q_a(0)}{1 - (\text{один сосед } f)} =$

$= \dfrac{g_{af}(0,0) - \varphi_{fa}(0) + q_a(0)}{1} = \dfrac{\mathbb{1}(0=0) - \overset{=0}{\varphi_f(0)} + q(0 = \overset{x_a}{0})}{1}$

$= \dfrac{1+1}{1} = 2.$

3) $\varphi_{af}(0) = g_{af}(0, k^*(f)) - \varphi_{fa}(k^*(f)) - C(a) =$

$= \mathbb{1}(0=0) - 0 - 2 = -1.$

Для „a". 1 — для метки 1 в объекте „a"

1) $k^*(f) \in \underset{k_{t'} \in \{0,1\}}{\operatorname{argmax}} \big(g_{af}(\underset{k_t}{1}, \underset{k_{t'}}{0}) - \varphi_{af}(1) - \varphi_{fa}(0),$

$g_{af}(1,1) - \varphi_{af}(1) - \varphi_{fa}(1) \big) \overset{\text{argmax}}{=} (0, 1) = \{1\}.$

2) $c(a) = \dfrac{g_{af}(1, k^*(f)) - \varphi_{fa}(k^*(f)) - q_a(1)}{1}$

3) $\varphi_{af}(1) = g_{af}(1, k^*(f)) - \varphi_{fa}(k^*(f)) - c(a) =$

$= 1 - 0 - 1 = 0.$

Аналогічне рішення буде і для об'єкта „в", бо інтерес такий самий як і в „а" та тей самий (єдиний) сусід „f".

$\Rightarrow \varphi_{bf}(0) = -1, \ \varphi_{bf}(1) = 0.$

Считаем для „d".

Соседи „d": „c" и „f"

Для d.0: , $k_t = 0$, $t = d$, $t' = \{„c", „f"\}$.

1) $k^*(„f") \in \underset{k_{t'} \in \{0,1\}}{\mathrm{argmax}} \left(g_{df}(0,0) - \varphi_{df}(0) - \varphi_{fd}(0), g_{df}(0,1) - \varphi_{df}(0) - \varphi_{fd}(1) \right) = \{0\}$

$k^*(„c") \in \underset{k_{t'} \in \{0,1\}}{\mathrm{argmax}} \left(g_{dc}(0,0) - \varphi_{dc}(0) - \varphi_{cd}(0), g_{dc}(0,1) - \varphi_{dc}(0) - \varphi_{cd}(1) \right) = \{0\}$

2) $c(d) = \dfrac{g_{df}(0, k^*(f)) + g_{dc}(0, k^*(c)) - \varphi_{fd}(k^*(f)) - \varphi_{cd}(k^*(c)) + q_d(0)}{2} = \dfrac{1 + 1 - 0 - 0 + 11(0=1)}{2} = 1.$

3) $\varphi_{df}(0) = g_{df}(0, k^*(f)) - \varphi_{fd}(k^*(f)) - c(d) = 1 - 0 - 1 = 0$

$\varphi_{dc}(0) = g_{dc}(0, k^*(c)) - \varphi_{cd}(k^*(c)) - c(d) = 1 - 0 - 1 = 0.$

Для d.1: , $k_t = 1$, $t = d$, $t' = \{„c", „f"\}$.

1) $k^*(f) \in \underset{k_{t'} \in \{0,1\}}{\mathrm{argmax}} \left(g_{df}(1,0) - \varphi_{df}(1) - \varphi_{fd}(0), g_{df}(1,1) - \varphi_{df}(1) - \varphi_{fd}(1) \right) = \{1\}$

$k^*(c) \in \underset{k_{t'} \in \{0,1\}}{\mathrm{argmax}} \left(g_{dc}(1,0) - \varphi_{dc}(1) - \varphi_{fc}(0), g_{dc}(1,1) - \varphi_{dc}(1) - \varphi_{cd}(1) \right) = \{1\}$

Scanned with CamScanner

$$= \frac{g_{df}(1, k^*(f)) + g_{dc}(1, k^*(c)) - \varphi_{fd}(k^*(f)) - \varphi_{cd}(k^*(c)) - g_d(1, \ldots}{2}$$

$$\frac{1 + 1 - 0 - 0 + 1}{2} = \frac{3}{2} = 1{,}5.$$

$$\varphi_{df}(1) = g_{df}(1, k^*(f)) - \varphi_{fd}(k^*(f)) - c(d) = 1 - 1{,}5 = -0{,}5$$

$$\varphi_{dc}(1) = g_{dc}(1, k^*(c)) - \varphi_{cd}(k^*(c)) - c(d) = 1 - 1{,}5 = -0{,}5.$$

Для підрахунку „f" (оскільки в нього сусіди a, b, c, d) випишемо φ:

$$\boxed{\begin{array}{ll} \varphi_{af}(0) = -1, & \varphi_{af}(1) = 0. \\ \varphi_{bf}(0) = -1, & \varphi_{bf}(1) = 0. \\ \varphi_{cf}(0) = 0, & \varphi_{cf}(1) = -0{,}5. \\ \varphi_{df}(0) = 0, & \varphi_{df}(1) = -0{,}5. \end{array}}$$

$$\varphi_{dc}(0) = 0, \quad \varphi_{dc}(1) = -0{,}5.$$

$$\varphi_{cd}(0) = 0, \quad \varphi_{cd}(1) = -0{,}5.$$

Рахуємо для f.:

$t = f$, $t' = \{a, b, c, d\}$.

Для f. 0.

$$k^*(a) = \underset{k_{t'} \in \{0,1\}}{\arg\max} \big(g_{fa}(0,0) - \varphi_{fa}(0) - \varphi_{af}(0), \; g_{fa}(0,1) - \varphi_{fa}(0) - \varphi_{af}(1)\big) = (1+1; 1) = \{0\}.$$

$$k^*(b) = \underset{k_{t'} \in \{0,1\}}{\arg\max} \big(g_{fb}(0,0) - \varphi_{fb}(0) - \varphi_{bf}(0), \; g_{fb}(0,1) - \varphi_{fb}(0) - \varphi_{bf}(1)\big) = (1+1; 0) = \{0\}.$$

$$k^*(c) = \underset{k_{t'} \in \{0,1\}}{\arg\max} \big(g_{fc}(0,0) - \varphi_{fc}(0) - \varphi_{cf}(0), \; g_{fc}(0,1) - \varphi_{fc}(0) - \varphi_{cf}(1)\big) = (1; +0{,}5) = \{0\}.$$

$$k^*(d) = \underset{k_{t'} \in \{0,1\}}{\arg\max} \big(g_{fd}(0,0) - \varphi_{fd}(0) - \varphi_{df}(0), \; g_{fd}(0,1) - \varphi_{fd}(0) - \varphi_{df}(1)\big) = (1; 0{,}5) = \{0\}.$$

$$C(f) = \frac{g_{fa}(0,k^*(a)) + g_{fb}(0,k^*(b)) + g_{fc}(0,k^*(c)) + g_{fd}(0,k^*(d)) - \varphi_{af}(k^*(a)) - \varphi_{bf}(k^*(b)) - \varphi_{cf}(k^*(c)) - \varphi_{df}(k^*(d)) + g_f(0)}{4} = \frac{1+1+1+1-\varphi_{af}(0)-\varphi_{bf}(0)-\varphi_{cf}(0)-\varphi_{df}(0)}{4} = \frac{4-(-1)-(-1)-0-0}{4} = \frac{6}{4} = \frac{3}{2} = 1{,}5.$$

$$\varphi_{fa}(0) = g_{fa}(0,k^*(a)) - \varphi_{af}(k^*(a)) - C(f) = 1+1-1{,}5 = 0{,}5.$$

$$\varphi_{fb}(0) = g_{fb}(0,k^*(b)) - \varphi_{bf}(k^*(b)) - C(f) = 1+1-1{,}5 = 0{,}5$$

$$\varphi_{fc}(0) = g_{fc}(0,k^*(c)) - \varphi_{cf}(k^*(c)) - C(f) = 1-0-1{,}5 = -0{,}5.$$

$$\varphi_{fd}(0) = g_{fd}(0,k^*(d)) - \varphi_{df}(k^*(d)) - C(f) = 1-0-1{,}5 = -0{,}5.$$

Для f-1

$$k^*(a) = \operatorname*{argmax}_{k_{f'}\in\{0,1\}} \big(g_{fa}(1,0) - \varphi_{fa}(1) - \varphi_{af}(0),\; g_{fa}(1,1) - \varphi_{fa}(1) - \varphi_{af}(1)\big) = (0+1, 1) = \{0\}.$$

$$k^*(b) = \operatorname*{argmax}_{k_{f'}\in\{0,1\}} \big(g_{fb}(1,0) - \varphi_{fb}(1) - \varphi_{bf}(0),\; g_{fb}(1,1) - \varphi_{fb}(1) - \varphi_{bf}(1)\big) = \{1,1\} = \{0\}.$$

$$k^*(c) = \operatorname*{argmax}_{k_{f'}\in\{0,1\}} \big(g_{fc}(1,0) - \varphi_{fc}(1) - \varphi_{cf}(0),\; g_{fc}(1,1) - \varphi_{fc}(1) - \varphi_{cf}(1)\big) = \{0, 1+0{,}5\} = \{1\}.$$

$$k^*(d) = 1 \ldots 1 = \{1\}.$$

k

$$\frac{g_{fa}(1, \overset{''0}{k^*(a)}) + g_{fb}(1, \overset{''0}{k^*(b)}) + g_{fc}(1, \overset{''1}{k^*(c)}) + g_{fd}(1, \overset{''1}{k^*(d)}) - \varphi_{af}(k^*(a)) - \varphi_{bf}(k^*(b)) - \varphi_{cf}(k^*(c)) - \varphi_{df}(k^*(d)) + g_f(1)}{4} = \frac{2+1+1+0{,}5+0{,}5}{4} = \frac{5}{4}$$

$$\varphi_{fa}(1) = g_{fa}(1, k^*(a)) - \varphi_{af}(k^*(a)) - C(f) = 1 - \frac{5}{4} = -\frac{1}{4}.$$

$$\varphi_{fb}(1) = g_{fb}(1, k^*(b)) - \varphi_{bf}(k^*(b)) - C(f) = 1 - \frac{5}{4} = -\frac{1}{4}.$$

$$\varphi_{fc}(1) = g_{fc}(1, k^*(c)) - \varphi_{cf}(k^*(c)) - C(f) = 1 + 0{,}5 - \frac{5}{4} = \frac{1}{4}.$$

$$\varphi_{fd}(1) = \frac{1}{4}.$$

Знайдемо розмітку

Для a: $a \downarrow f$

любой сосед для a: $\arg\max$

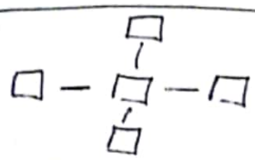
$$\begin{cases} \max\begin{cases} g_{af}(\overset{k\ k'}{0,0}) - \varphi_{af}(\overset{k}{0}) - \varphi_{fa}(\overset{k'}{0}) = 1{,}5 \\ g_{af}(0,1) - \varphi_{af}(0) - \varphi_{fa}(1) = \frac{5}{4} \end{cases} \xrightarrow{\max} 1{,}5 \\ \max\begin{cases} g_{af}(1,0) - \varphi_{af}(1) - \varphi_{fa}(0) = -0{,}5 \\ g_{af}(1,1) - \varphi_{af}(1) - \varphi_{fa}(1) = \frac{5}{4} \end{cases} \xrightarrow{\max} \frac{5}{4} \end{cases} \xrightarrow{\arg\max} = 0.$$

Для $f \downarrow b$: $k = 0$. , Для $c - f$: $k = 0$, Для $d \downarrow c$ $k = 1$.

Для $b \downarrow f$ $k = 0$.

Вивести частковий алгоритм (випадок) TRW-S - для покращення оцінки зверху ймовірності найбільш ймовірного стану вип. поле Маркова на графі-решітці.

Структура сусідства : $|N(t)| \le 4$:



$T = \{1, \dots, m\} \times \{1, \dots, n\}$

$\tau = \{(i,j), (i+1,j) \mid i = \overline{1,m-1}, j = \overline{1,n}\} \cup \{(i,j), (i,j+1) \mid i = \overline{1,m}, j = \overline{1,n-1}\}$

Розділимо розмітку на горизонтальну (Г) та вертикальну (В).

Введемо ф-ії:

$\alpha^{Г}(\bar{\ell})$ - задає вагу розмітки горизонтальної i; $\bar{\ell} \in K^n$ - розмітка рядка чи стовпця.

$\alpha^{Г} : K^n \to \mathbb{R}$; $\alpha^{Г}(\bar{\ell}) \ge 0$; $\sum\limits_{\ell \in K^n} \alpha^{Г} = 1.$

Оскільки ми ділимо розмітку то і вертикальній і в горизонтальній будуть $q_{ij}(k) \Rightarrow$ розділимо $q_{ij}(k)$ наприклад на $\frac{1}{2}$

$\frac{1}{2} q_{ij}(k) + \frac{1}{2} q_{ij}(k)$

Введемо : $Q_i^{Г} : K^n \to \mathbb{R}$ - якість розмітки (горизонтальної) $\bar{\ell}$

Аналогічно для $Q_j^{В}$

Визначимо : $Q_i^{Г}(\bar{\ell}) = \sum\limits_{j=1}^{n} \frac{1}{2} q_{ij}(\ell_j) + \sum\limits_{j=1}^{n-1} g_{ij\,ij+1}(\ell_i, \ell_{j+1})$; $i = \overline{1,m}$, $\bar{\ell} \in K^n$.

Аналогічно для стовпця:

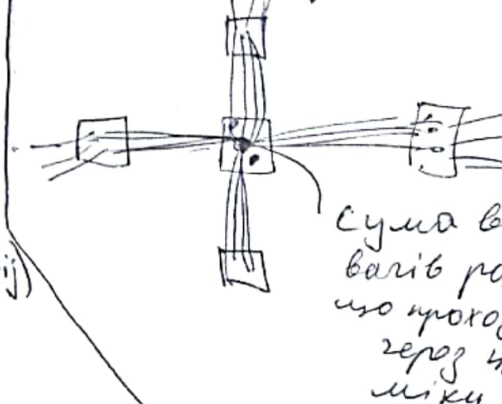
$$Q_j^{В}(\bar{\ell}) = \sum_{i=1}^{m} \frac{1}{2} q_{ij}(\ell_i) + \sum_{i=1}^{m-1} g_{ij\,i+1j}(\ell_i, \ell_{i+1}) \;;\; j = \overline{1,n},\; \bar{\ell} \in K^m.$$

Запишемо пряму задачу:

$$\sum_{i=1}^{m} \sum_{\bar{\ell} \in K^n} \alpha_i^{Г}(\bar{\ell}) \cdot Q_i^{Г}(\bar{\ell}) + \sum_{j=1}^{n} \sum_{\bar{\ell} \in K^m} \alpha_j^{В}(\bar{\ell}) \cdot Q_j^{В}(\bar{\ell}) \xrightarrow{\alpha} \max.$$

$$\begin{cases} \sum\limits_{\bar{\ell} \in K^n} \alpha_i^{Г}(\bar{\ell}) = 1, \; i = \overline{1,m} & \Big| \; \theta_i \\ \sum\limits_{\bar{\ell} \in K^m} \alpha_j^{В}(\bar{\ell}) = 1, \; j = \overline{1,m} & \Big| \; \gamma_j \\ \sum\limits_{\substack{\bar{\ell} \in K^n \\ \ell_j = k_{ij}}} \alpha_i^{Г}(\bar{\ell}) - \sum\limits_{\substack{\bar{\ell} \in K^m \\ \ell_i = k_{ij}}} \alpha_j^{В}(\bar{\ell}) = 0 \;;\; i = \overline{1,m},\; j = \overline{1,n},\; k_{ij} \in K. & \Big| \; \varphi_{ij}(k_{ij}) \\ \alpha_i^{Г} \ge 0 \\ \alpha_j^{В} \ge 0 \end{cases}$$

Аналогічно дифузії:



Сума всіх гориз. вагів розміток що проходять через цю мітку має = сумі всіх верт. вагів розміток.

Запишемо двоїсту задачу:

$$\sum_{i=1}^{m} Q_i + \sum_{j=1}^{n} \gamma_j \xrightarrow{Q,\gamma} \min$$

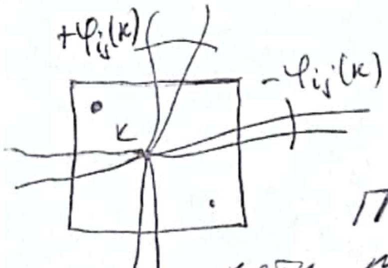
$$\lambda_i^{\Gamma}(\ell): Q_i + \sum_{j=1}^{n} \varphi_{ij}(\ell_j) \geq Q_i^{\Gamma}(\bar{\ell}), \; i=\overline{1,m}, \; \bar{\ell}\in K^n$$

$$\lambda_j^{B}(\ell): \gamma_j - \sum_{i=1}^{m} \varphi_{ij}(\ell_i) \geq Q_j^{B}(\bar{\ell}), \; j=\overline{1,n}, \; \bar{\ell}\in K^m$$

$$\Rightarrow Q_i = \max_{\bar{\ell}\in K^n}\left(Q_i^{\Gamma}(\bar{\ell}) - \sum_{j=1}^{n}\varphi_{ij}(\ell_j)\right); \; \gamma_j = \max_{\bar{\ell}\in K^m}\left(Q_j^{B}(\bar{\ell}) + \sum_{i=1}^{m}\varphi_{ij}(\ell_i)\right)$$

Підставимо в цільову ф-ю:

$$\sum_{i=1}^{m} \max_{\bar{\ell}\in K^n}\left(Q_i^{\Gamma}(\bar{\ell}) - \sum_{j=1}^{n}\varphi_{ij}(\ell_j)\right) + \sum_{j=1}^{n} \max_{\bar{\ell}\in K^m}\left(Q_j^{B}(\bar{\ell}) + \sum_{i=1}^{m}\varphi_{ij}(\ell_i)\right) \xrightarrow{\varphi} \min$$



Ми перекошимо якість $\varphi_{ij}(k)$ з горизонтальних розміток, що проходять через мітку k на вертикальні розмітки, що проходять через k. По аналогії з дифузією (там ми хотіли прирівняти max дуги) тут ми намагаємось прирівняти „максимальну" горизонтальну и „максимальну" гор. розмітку через k. об'єкта ij:

$$\underbrace{\max_{\substack{\bar{\ell}\in K^n \\ \ell_j=k}}\left(Q_i^{\Gamma}(\bar{\ell}) - \sum_{j=1}^{n}\varphi_{ij}(\ell_j)\right)} = \underbrace{\max_{\substack{\bar{\ell}\in K^m \\ \ell_i=k}}\left(Q_j^{B}(\bar{\ell}) + \sum_{i=1}^{m}\varphi_{ij}(\ell_i)\right)}$$

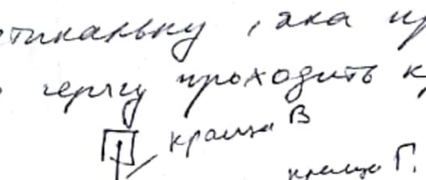
– Таким чином ми мінімізуємо цільову ф-ю

Нехай $\Delta = (*) - (**) \Rightarrow \varphi_{ij}(k) += \frac{\Delta}{2}$

та розмітка, що більша ми її зменшуємо, навпаки

Нехай ми досягли, що краща верт. розмітка через k об'єкта ij = = кращій горизонт.

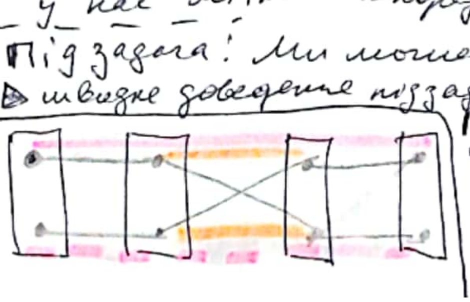
Покажемо, що, якщо задача супермодулярна, то ми можемо знайти таку кращу розмітку вертикальну, яка проходить через k об'єкта ij, через який в свою чергу проходить краща горизонтальна. Тобто ~~...~~ буде так:



Да, кращих розміток може бути декілька, але через супермодулярність у нас мітки впорядковані, тому ми беремо „найвищі" мітки

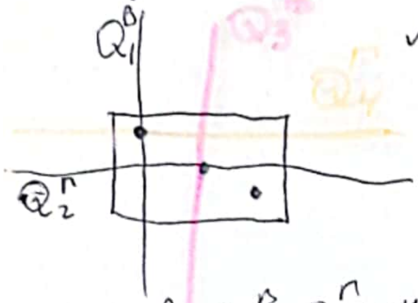
Підзадача: Ми можемо в кожному рядку вибрати максимальну розмітку

▲ швидке доведення підзадачі:



По озн. супер: сума двох — дуг більше за суму двох — то якщо — максимальні => — та — максимальні => можна однозначно вибрати ~~найвищу~~ найкращу розмітку ▲ (не повне дов-ння)

Доведемо



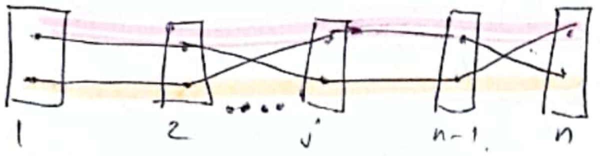
Ми знаємо, що $Q_1^B \geq Q_3^B$, - де це якщо, ми ~~допуст~~ припустили, що $Q_2^Г \geq Q_1^Г$

$$Q_1^B \geq Q_3^B = Q_2^Г \Rightarrow Q_1^B = Q_3^B$$
$$Q_2^Г \geq Q_4^Г = Q_1^B \qquad Q_2^Г = Q_4^Г$$

вони однакові => вибираємо наприклад $Q_1^B - Q_4^Г$

$Q_1^B, Q_2^Г, Q_3^B, Q_4^Г$ - максимальні (припускаємо)

Нехай $\bar{k}$ і $\bar{\ell}$ - максимальні розмітки:



Доведемо, що — і — будуть максимальними:

$\bar{k} \in K^n, k_j$; $k_j \vee k_j$

$\bar{\ell} \in K^n, \ell_j$; $k_j \wedge k_j$

Для кожної пари об'єктів розглянемо із супермодулярності:

$$g_{jj+1}(k_j, k_{j+1}) + g_{jj+1}(\ell_j, \ell_{j+1}) \leq g_{jj+1}(k_j \vee \ell_j, k_{j+1} \vee \ell_{j+1}) +$$
$$+ g_{jj+1}(k_j \wedge \ell_j, k_{j+1} \wedge \ell_{j+1}), \; j = \overline{1, n-1}.$$

Виразимо через якості розміток:

$$\underbrace{\sum_{j=1}^{n-1} g_{jj+1}(k_j, k_{j+1})}_{Q(\bar{k})} + \underbrace{\sum_{j=1}^{n-1} g_{jj+1}(\ell_j, \ell_{j+1})}_{Q(\bar{\ell})} \leq \underbrace{\sum_{j=1}^{n-1} g_{jj+1}(k_j \vee \ell_j, k_{j+1} \vee \ell_{j+1})}_{Q(\bar{k} \vee \bar{\ell})} +$$
$$+ \underbrace{\sum_{j=1}^{n-1} g_{jj+1}(k_j \wedge \ell_j, k_{j+1} \wedge \ell_{j+1})}_{Q(\bar{k} \wedge \bar{\ell})}.$$

Q - поелементний (максимум/мінімум).

$$\begin{cases} Q(\bar{k}) + Q(\bar{\ell}) \leq Q(\bar{k} \vee \bar{\ell}) + Q(\bar{k} \wedge \bar{\ell}). \\ Q(\bar{k}) \geq Q(\bar{k} \vee \bar{\ell}) ; \; Q(k) \geq Q(\bar{k} \wedge \bar{\ell}) \\ Q(\bar{\ell}) \geq Q(\bar{k} \wedge \bar{\ell}) ; \; Q(k) \geq Q(\bar{k} \vee \bar{\ell}) \end{cases} \Rightarrow Q(\bar{k}) = Q(\bar{\ell}) = Q(\bar{k} \vee \bar{\ell}) = Q(\bar{k} \wedge \bar{\ell})$$

=> Вибір найвищої максимальної розмітки однозначний!

Введемо:

$$L_{ij}(k) = \max_{\substack{\bar{\ell} \in K^{j-1} \\ \ell_j = k}} \left[\sum_{t=1}^{j-1} \left(\frac{1}{2} g_{is}(\ell_i) - \varphi_{it}(\ell_t) + g_{it\,it+1}(\ell_t, \ell_{t+1}) \right) \right]$$

півмене K^{j-1}

$$R_{ij}(k) = \max_{\substack{\bar{\ell} \in K^{n-j} \\ \ell_j = k}} \left[\sum_{t=j+1}^{n} \left(\frac{1}{2} g_{it}(\ell_t) - \varphi_{it}(\ell_t) + g_{it-1\,it}(\ell_{t-1}, \ell_t) \right) \right]$$

Якість найкращої розмітки зліва для i,j

аналогічно для R

Scanned with CamScanner

$$\max_{\substack{\bar\ell\in K^n\\ \ell_j=k}}\left(Q_i^{\Gamma}(\bar\ell)-\sum_{j=1}^{n}\varphi_{ij}(\ell_j)\right)=L_{ij}(k)+R_{ij}(k)+\frac12 q_{ij}(k)-\varphi_{ij}(k)$$

Як рахувати $L_{ij+1}(k)$, $R_{ij-1}(k)$ – ? → через дин. програмування

$$L_{ij+1}(k)=\max_{k'\in K}\left[L_{ij}(k')+\frac12 q_{ij}(k')-\varphi_{ij}(k')+g_{ijij+1}(k',k)\right]$$

$$R_{ij-1}(k)=\max_{k'\in K}\left[R_{ij}(k')+\frac12 q_{ij}(k')\right.$$

Важливо: $R_{in}(k)=0$, $L_{i1}(k)=0$

Після репараметризації:

Аналогічно для U, D.

Ми знаємо: $$\max_{\substack{\bar\ell\in K^n\\ \ell_j=k}}\left(Q_i^{\Gamma}(\bar\ell)-\sum_{j=1}^{n}\varphi_{ij}(\ell_j)\right)=L_{ij}(k)+R_{ij}(k)+\frac12 q_{ij}(k)-\varphi_{ij}(k)$$

$$\max_{\substack{\bar\ell\in K^m\\ \ell_i=k}}\left(Q_j^{B}(\bar\ell)+\sum_{i=1}^{m}\varphi_{ij}(\ell_i)\right)=U_{ij}(k)+D_{ij}(k)+\frac12 q_{ij}(k)+\varphi_{ij}(k).$$

$$\Delta=L_{ij}(k)+R_{ij}(k)+\cancel{\frac12 q_{ij}(k)}-\varphi_{ij}(k)-U_{ij}(k)-D_{ij}(k)-\cancel{\frac12 q_{ij}(k)}-\varphi_{ij}(k)$$

$$\Rightarrow\frac{\Delta}{2}=\left(L_{ij}(k)+R_{ij}(k)-U_{ij}(k)-D_{ij}(k)/2\right)-\varphi_{ij}(k)$$

Ми пам'ятаємо: $\varphi_{ij}\pm\frac{\Delta}{2}$

Алгоритм TRW-S (частковий)

Ініціалізація: $R=U, D=0$, $U_{1j}=L_{i1}=D_{nj}=R_{im}=0$

for $i=\overline{1,m}$:
for $j=\overline{1,n}$:
for $k\in K$:

$$D_{ij}(k)=\max_{k'\in K}\left[D_{i+1j}(k')+\frac12 q_{i+1j}(k')+g_{iji+1j}(k,k')\right]$$

$$R_{ij}(k)=\max_{k'\in K}\left[R_{ij+1}(k')+\frac12 q_{ij+1}(k')+g_{ijij+1}(k,k')\right]$$

Прямий прохід: for $i=\overline{1,m}$
for $j=\overline{1,n}$
for $k\in K$

$$L_{ij}(k)=\max_{k'\in K}\left[L_{ij-1}(k')+\frac12 q_{ij-1}(k')-\varphi_{ij-1}(k')+g_{ij-1ij}(k',k)\right]$$

$$U_{ij}(k)=\max_{k'\in K}\left[U_{i-1j}(k')+\frac12 q_{i-1j}(k')+\varphi_{i-1j}(k')+g_{i-1jij}(k',k)\right]$$

$$\varphi_{ij}(k)=\left(L_{ij}(k)+R_{ij}(k)-U_{ij}(k)-D_{ij}(k)\right)/2.$$

Зворотній прохід: for $i=\overline{m,1}$:
for $j=\overline{n,1}$:
for $k\in K$:

$$D_{ij}(k)=\max_{k'\in K}\left[D_{i+1j}(k')+\frac12 q_{i+1j}(k')+\varphi_{i+1j}(k')+g_{iji+1j}(k,k')\right]$$

$$R_{ij}(k)=\max_{k'\in K}\left[R_{ij+1}(k')+\frac12 q_{ij+1}(k')-\varphi_{ij+1}(k')+g_{ijij+1}(k,k')\right]$$

$$\varphi_{ij}(k)=\left(L_{ij}(k)+R_{ij}(k)-U_{ij}(k)-D_{ij}(k)\right)/2$$

Розмітка: $\forall ij$: $k^*_{ij}=\arg\max_{k'\in K}\left(L_{ij}(k')+R_{ij}(k')+\frac12 q_{ij}(k')-\varphi_{ij}(k')\right).$

Маємо зображення $x : T \to [0;1]^3$ та його часткову сегментацію $s^0 : T \to \{f, g, \varepsilon\}$, де ε – відсутня інформація. Вказати алгоритм, що ітеративно покращує сегментацію s, якщо відомі $N^f, N^b \in \mathbb{N}, \gamma \in \mathbb{R}^+, \beta$ та

$$p(s;x) = \frac{1}{Z(x)} \cdot \prod_{t \in T} q_t(s_t; x) \cdot \prod_{tt' \in T} \exp\left\{ -\mathbb{I}(s_t \neq s_{t'}) \cdot \gamma \cdot e^{-\frac{\|x_t - x_{t'}\|^2}{2\beta}} \right\}$$

$$q_t(s_t; x) = \sum_{i=1}^{N^{s_t}} p^{s_t}(i) \cdot (2\pi)^{-\frac{5}{2}} \cdot \det(\Sigma_i^{s_t})^{-\frac{1}{2}} \cdot \exp\left\{ -\frac{1}{2} (x_t - \mu_i^{s_t})^T \cdot (\Sigma_i^{s_t})^{-1} (x_t - \mu_i^{s_t}) \right\}$$

Випишемо TRW-S та GMM.

TRW-S

Ініціалізація: $R \equiv 0, D \equiv 0;\ U_{1j} = L_{i1} \equiv D_{nj} = R_{im} = 0$

for $i = \overline{1,m}$:
 for $j = \overline{1,n}$:
  for $k \in K$:

$$D_{ij}(k) = \max_{k' \in K} \left[D_{i+1j}(k') + \frac{1}{2} q_{i+1j}(k') + g_{ij\,i+1j}(k, k') \right]$$

$$R_{ij}(k) = \max_{k' \in K} \left[R_{ij+1}(k') + \frac{1}{2} q_{ij+1}(k') + g_{ij\,ij+1}(k, k') \right]$$

Прямий прохід:

for $i = \overline{1,m}$:
 for $j = \overline{1,n}$:
  for $k \in K$:

$$L_{ij}(k) = \max_{k' \in K} \left[L_{ij-1}(k') + \frac{1}{2} q_{ij-1}(k') - \varphi_{ij-1}(k') + g_{ij-1\,ij}(k', k) \right].$$

$$U_{ij}(k) = \max_{k' \in K} \left[U_{i-1j}(k') + \frac{1}{2} q_{i-1j}(k') + \varphi_{i-1j}(k') + g_{i-1j\,ij}(k', k) \right].$$

$$\varphi_{ij}(k) = (L_{ij}(k) + R_{ij}(k) - U_{ij}(k) - D_{ij}(k))/2$$

Зворотній прохід:

for $i = \overline{m,1}$:
for $j = \overline{n,1}$:
for $k \in K$:

$$D_{ij}(k) = \max_{k' \in K} \left[D_{i+1j}(k') + \frac{1}{2} q_{i+1j}(k') + \varphi_{i+1j}(k') + g_{ij\,i+1j}(k, k') \right]$$

$$R_{ij}(k) = \max_{k' \in K} \left[R_{ij+1}(k') + \frac{1}{2} q_{ij+1}(k') - \varphi_{ij+1}(k') + g_{ij\,ij+1}(k, k') \right]$$

$$\varphi_{ij}(k) = (L_{ij}(k) + R_{ij}(k) - U_{ij}(k) - D_{ij}(k))/2.$$

Розмітка:

$$\forall i,j:\ k^*_{ij} = \arg\max_{k' \in K} \left(L_{ij}(k') + R_{ij}(k') + \frac{1}{2} q_{ij}(k') - \varphi_{ij}(k') \right)$$

EM для GMM

E step

рахуємо $p(k|x)$.

M step

$$\Sigma_k = \frac{\sum_{i=1}^{n} p(k|x) \cdot (x - \mu)(x - \mu)^T}{\sum_{i=1}^{n} p(k|x)}, \quad \mu_k = \frac{\sum_{i=1}^{n} p(k|x_i, \mu_k, \Sigma_k) \cdot x_i}{\sum_{i=1}^{n} p(k|x_i, \mu_k, \Sigma_k)}$$

Алгоритм

1. Беремо всі пікселі як датасет, що належать foreground і background.

2. Подаємо на вхід дані пікселі для EM алгоритму.
Отримуємо набори.
$\{\langle p^f(i), \mu_i^f, \Sigma_i^f \rangle : i = \overline{1, N^f}\}$ та $\{\langle p^b(i), \mu_i^b, \Sigma_i^b \rangle : i = \overline{1, N^b}\}$
Рахуємо $q_t(s_t; x)$ та $g_{tt'}(s_t, s_{t'}, x)$

3. Подаємо на вхід q та g до TRW-S
Отримуємо розмітку.
Оновлюємо датасет по новій розмітці.

Повторюємо кроки 1-3 n_iter кількість разів.

Для випадкового поля Маркова з розподілом ймовірностей $p(\bar{k}) = \frac{1}{Z} \cdot \exp\{\sum_{tt' \in \tau} g_{tt'}(k_t, k_{t'})\}$, $\forall \bar{k}: T \to K$ показати, якій умові має задовольняти максимально правдоподібна оцінка невідомих параметрів $g_{tt'}$ по навчальній вибірці $\bar{k}^1, \bar{k}^2, \bar{k}^3 \ldots \bar{k}^n \in K^T$? Як зміниться відповідь, якщо відомо, що всі $g_{tt'}$ однакові для всіх $tt' \in \tau$?

Випадок 1. Нехай всі функції між об'єктами однакові.

Вважаємо, що всі $g > 0$

Z-нормуючий коефіцієнт. $\sum_{\bar{k} \in K^T} p(\bar{k}) = 1 \Leftrightarrow Z(g) = \sum_{\bar{k} \in K^T} \exp \sum_{tt' \in \tau} g(k_t, k_{t'})$

Яке має бути g?

$$l(g) = \prod_{i=1}^{n} p(\bar{k}^i) \leftarrow \text{ф-я правдоподібності.}$$

Нехай $L(g) = \log l(g) = \sum_{i=1}^{n} \log p(\bar{k}^i)$

$$\Rightarrow g = \arg\max_{g^*} l(g^*) = \{\log \text{ - монот. ф-я} \uparrow\} \arg\max_{g^*} L(g^*)$$

$$L(g) = \sum_{i=1}^{n} \log\left[\frac{1}{Z(g)} \cdot \exp \sum_{tt' \in \tau} g(k_t^i, k_{t'}^i)\right] = \left\{\begin{matrix}\text{Візьмемо } \log \\ \text{від добутку}\end{matrix}\right\} =$$

$$= \sum_{i=1}^{n} \sum_{tt' \in \tau} g(k_t^i, k_{t'}^i) - n \log Z(g).$$

$$n(k_1, k_2) = \sum_{i=1}^{n} \sum_{tt' \in \tau} \mathbb{1}(k_t^i = k_1 \; \& \; k_{t'}^i = k_2) \text{ - скільки разів зустрічається пара}$$

Перепишемо $L(g)$.

$$L(g) = \sum_{k_1 \in K} \sum_{k_2 \in K} n(k_1, k_2) \cdot g(k_1, k_2) - n \log Z(g)$$

$$\frac{\partial L}{\partial g(x,y)} = n(x,y) - n \cdot \frac{1}{Z(g)} \cdot \frac{\partial Z}{\partial g(x,y)} = 0.$$

$$Z(g) = \sum_{\bar{k} \in K^T} \exp \sum_{tt' \in \tau} g(k_t, k_{t'})$$

$$\frac{\partial Z}{\partial g(x,y)} = \sum_{\bar{k} \in K^T} \exp \sum_{tt' \in \tau} g(k_t, k_{t'}) \cdot \frac{\partial \sum_{tt' \in \tau} g(k_t, k_{t'})}{\partial g(x,y)} =$$

в цій сумі має зустрітись $g(x,y)$

$$= \sum_{\bar{k} \in K^T} \exp\left[\sum_{tt' \in \tau} g(k_t, k_{t'})\right] \sum_{tt' \in \tau} \frac{\partial g(k_t, k_{t'})}{\partial g(x,y)} =$$

$$= \sum_{\bar{k} \in K^T} \exp\left[\sum_{tt' \in \tau} g(k_t, k_{t'})\right] n_{\bar{k}}(x,y)$$

$$\frac{n(x,y)}{n} = \frac{1}{Z(g)} \sum_{\bar{k} \in K^T} n_{\bar{k}}(x,y) \cdot \exp \sum_{tt' \in \tau} g(k_t, k_{t'}) = \left\{\begin{matrix}\text{внесемо} \\ \frac{1}{Z(g)}\end{matrix}\right\}$$

$$\Rightarrow \boxed{\underbrace{\frac{n(x,y)}{n}}_{\text{оцінка мат очікування}} = \underbrace{\sum_{\bar{k} \in K^T} n_{\bar{k}}(x,y) \cdot p(\bar{k})}_{\text{мат. очікування пари } (x,y) \text{ в розмітці}}}$$

Випадок 2. Між всіма парами об'єктів у нас різні ф-ії.

$$p(k) = \frac{1}{Z(g)} \exp \sum_{tt' \in \tau} g_{tt'}(k_t, k_{t'})$$

$$Z(g) = \sum_{k \in K^T} \exp \sum_{tt' \in \tau} g_{tt'}(k_t, k_{t'}) ; \quad L(g) = \sum_{i=1}^{n} \log \frac{1}{Z(g)} \exp \sum_{tt' \in \tau} g_{tt'}(k_t, k_{t'}) =$$

$$= \sum_{i=1}^{n} \sum_{tt' \in \tau} g_{tt'}(k_t, k_{t'}) - n \cdot Z(g)$$

$$n_{tt'}(k_1, k_2) = \sum_{i=1}^{n} \mathbb{1}(k_t^i = k_1 \ \& \ k_{t'}^i = k_2); \quad L(g) = \sum_{tt' \in \tau} \sum_{k_1 \in K} \sum_{k_2 \in K} n_{tt'}(k_1, k_2) \cdot$$

$$\cdot g_{tt'}(k_1, k_2) - n \cdot (Z(g));$$

$$\frac{\partial L}{\partial g_{tt'}(x,y)} = n_{tt'}(x,y) - n \cdot \frac{\frac{\partial Z}{\partial g_{tt'}(x,y)}}{Z(g)} = 0.$$

$$\frac{\partial Z}{\partial g_{tt'}(x,y)} = \frac{\partial}{\partial g_{tt'}(x,y)} \left(\sum_{k \in K^T} \exp \sum_{tt' \in \tau} g_{tt'}(k_t, k_{t'}) \right) = \sum_{k \in K^T} \exp \left[\sum_{tt' \in \tau} g_{tt'}(k_t, k_{t'}) \right] \cdot$$

$$\cdot \frac{\partial \sum_{ss' \in \tau} g_{ss'}(k_t, k_{t'})}{\partial g_{tt'}(x,y)} = \sum_{k \in K^T} \exp \left[\sum_{tt' \in \tau} g_{tt'}(k_t, k_{t'}) \right] \sum_{ss' \in \tau} \frac{\partial g_{ss'}(k_s, k_{s'})}{\partial g_{tt'}(x,y)} =$$

позначили через ss'

$$= |-| \frac{\partial g_{tt'}(k_t, k_{t'})}{\partial g_{tt'}(x,y)} = \sum_{\substack{k \in K^T \\ k_t = x \\ k_{t'} = y}} \exp \left[\sum_{tt' \in \tau} g_{tt'}(k_t, k_{t'}) \right]$$

$$n_{tt'}(x,y) - \frac{n}{Z(g)} \cdot \frac{\partial Z}{\partial g_{tt'}(x,y)} = 0 \Rightarrow \frac{n_{tt'}(x,y)}{n} = \sum_{\substack{k \in K^T \\ k_t = x \\ k_{t'} = y}} \underbrace{\frac{1}{Z(g)} \exp \left[\sum_{tt' \in \tau} g_{tt'}(k_t, k_{t'}) \right]}_{p(\bar{k})}$$

$$\Rightarrow \underbrace{\frac{n_{tt'}(x,y)}{n}}_{\text{частота зустрічі пари } (x,y)} = \sum_{\substack{\bar{k} \in K^T \\ k_t = x \\ k_{t'} = y}} p(\bar{k}) = P_{tt'}(x,y)$$

ймовірність того, що у вибірці зустрілась пара (x,y)

EM - алгоритм для покращення максимально правдоподібної оцінки. Доведення монотонності.

Введемо позначення:

$\hat{X} = \bar{x}_1, \dots, \bar{x}_n$; $\hat{k} = \bar{k}_1, \dots, \bar{k}_n$; $\bar{x}_i \in X^T, \bar{k}_i \in K^T$; $\hat{x} \in \hat{X} = (X^T)^n$; $\hat{k} \in \hat{K}$

Сумісна ймовірність з параметром θ

$$p(\hat{x}, \hat{k}, \theta) = \prod_{i=1}^{n} p(\bar{x}_i, \bar{k}_i, \theta)$$

маргінальна ймовірність

Головна ціль: $L(\theta) = \log p(\hat{x}; \theta) = \log \sum_{k} p(\bar{x}, \bar{k}; \theta) \xrightarrow{\theta} max$

$p(\hat{x}; \theta^*) = \{$розпишемо через умовну ймов-ть$\} = \dfrac{p(\hat{x}, \hat{k}, \theta^*)}{p(\hat{k}/\hat{x}, \theta^*)}$ $\forall$ $k \in K$. (*)

$\Rightarrow$ Тоді $L(\theta^*) = \log \dfrac{p(\hat{x}, \hat{k}, \theta^*)}{p(\hat{k}/\hat{x}, \theta^*)} = \left\{ \sum_{\hat{k} \in \hat{K}} p(\hat{k}/\hat{x}, \theta^j) = 1;\ \text{домножимо} \right\} =$

$= \left[\sum_{\hat{k} \in K} p(\hat{k}/\hat{x}; \theta^j) \right] \cdot \log \dfrac{p(\hat{x}, \hat{k}, \theta^*)}{p(\hat{k}/\hat{x}; \theta^*)} = \{$внесемо $\log$ під суму, бо (*)$\} \Rightarrow$

$$\Rightarrow L(\theta^*) = \sum_{\hat{k} \in K} p(\hat{k}/\hat{x}; \theta^j) \cdot \log p(\hat{x}, \hat{k}, \theta^*) \quad 1)$$

$$- \sum_{\hat{k} \in \hat{K}} p(\hat{k}/\hat{x}; \theta^j) \cdot \log p(\hat{k}/\hat{x}, \theta^*) \quad 2).$$

$$L(\theta^j) = \sum_{\hat{k} \in \hat{K}} p(\hat{k}/\hat{x}; \theta^j) \cdot \log p(\hat{x}, \hat{k}, \theta^j) \quad 3).$$

$$- \sum_{\hat{k} \in \hat{K}} p(\hat{k}/\hat{x}; \theta^j) \cdot \log p(\hat{k}/\hat{x}; \theta^j) \quad 4).$$

У 4) записана ентропія $\Rightarrow$ 4) $\leq$ 2). $\rightarrow$ за лемою Шеннона

$$\theta^{j+1} = \arg\max_{\theta^*} \sum_{\hat{k} \in K} p(\hat{k}/\hat{x}; \theta^j) \cdot \log p(\hat{x}, \hat{k}; \theta^*)$$

Якщо ми візьмемо таке θ^{j+1}, що буде максимізувати вираз

$\Rightarrow$ 3) $\geq$ 1. $\Rightarrow$ $L(\theta^{j+1}) \geq L(\theta^j)$.

Далі розпишемо

$\sum_{\hat{k} \in K} p(\hat{k}/\hat{x}; \theta^j) \cdot \log p(\hat{x}, \hat{k}; \theta^*) = \{$розпишемо покоординатно$\} =$

$= \sum_{\hat{k} \in K} \prod_{i=1}^{n} p(\bar{k}_i / \bar{x}_i; \theta^j) \cdot \log \prod_{s=1}^{n} p(\bar{x}_s, \bar{k}_s; \theta^*) = \{$розпишемо $\log$ добутку$\} =$

$$= \sum_{s=1}^{n} \sum_{\hat{k} \in \hat{K}} \prod_{i=1}^{n} p(\bar{k}_i / \bar{x}_i; \theta^j) \cdot \log p(\bar{x}_s, \bar{k}_s, \theta^*) =$$

$= \{$ розпишемо по кожному $k_i \} =$

$$= \sum_{s=1}^{n} : \begin{array}{l} \left[\sum_{k_1 \in K^T} p(\bar{k}_1 / \bar{x}_1; \theta^j) \right] = 1 \\ \dots \\ \left[\sum_{\bar{k}_s \in K^T} p(\bar{k}_s / \bar{x}_s; \theta^j) \cdot \log p(\bar{x}_s, \bar{k}_s, \theta^*) \right] \\ \dots \\ \left[\sum_{k_n \in K^T} p(\bar{k}_n / \bar{x}_n, \theta^j) \right] = 1 \end{array}$$

$$= \sum_{s=1}^{n} \sum_{\bar{k}_s \in K^T} p(\bar{k}_s / \bar{x}_s; \theta^j) \cdot \log p(\bar{x}_s, \bar{k}_s, \theta^*) =$$

$$= \sum_{i=1}^{n} \sum_{k \in K^T} p(\bar{k}_i / x_i; \theta^j) \cdot \log p(\bar{x}_i, \bar{k}_i; \theta^*).$$

Алгоритм

Якщо задано θ^m, то обрахування θ^{m+1}, можна розділити на 2 кроки.

1. E step

   Знаходимо $p(\bar{k} / \bar{x}; \theta^m)$ та $p(\bar{x}, \bar{k}, \theta^m)$.

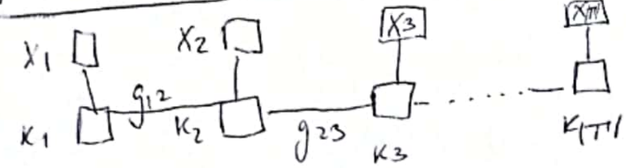
2. M step

   $$\theta^{m+1} = \underset{\theta^*}{\operatorname{argmax}} \sum_{\bar{k} \in K^n} p(\bar{k} / \bar{x}; \theta^m) \cdot \log p(\bar{x}, \bar{k}, \theta^*).$$

Нехай випадкове поле з прихованими змінними $\bar{k} \in K^T$ та змінними $\bar{x} \in \bar{X}^T$, що спостерігаються, має розподіл ймовірностей

$$p(\bar{k},\bar{x}) = \prod_{t\in T}\left(p(x_t \mid k_t)\cdot \frac{1}{Z}\exp\left\{\sum_{tt'\in T} g_{tt'}(k_t,k_{t'})\right\}\right).$$

Нехай параметри $g_{tt'}$ відомі для всіх $tt'\in T$. Показати, як застосувати EM-алгоритм для покращення макс. правдоподібної оцінки умовних ймовірностей $p(x|k)$ по навчальній вибірці $\bar{x}^1, \bar{x}^2, \bar{x}^3, \ldots \bar{x}^n \in X^T$



g визначають розподіл ймовірності на розмітці кожен із x залежить від свого k! Якщо ми знаємо k => ми знаємо розподіл $p(x_t/k_t;\theta_{k_t})$ з точністю до якогось θ_{k_t}.

$$p(\bar{x}/\bar{k};\bar{\theta}) = \prod_{t\in T} p(x_t/k_t;\theta_{k_t}) \quad (\blacktriangle)$$

Якщо ми знаємо $p(\bar{x},\bar{k},\bar{\theta})$ => ми знаємо і:

$$p(\bar{x},\bar{k};\bar{\theta}) = p(\bar{x}/\bar{k};\bar{\theta})\cdot p(\bar{k})$$

Ф-я правдоподібності: $l(\bar\theta) = \prod_{i=1}^{n} p_{\bar x}(\bar x^i;\bar\theta) = \prod_{i=1}^{n} \sum_{\bar k\in K^T} p(\bar x^i;\bar k;\bar\theta) \xrightarrow{\bar\theta} \max$

Згадаємо EM алгоритм: $\bar\theta^1 \to \bar\theta^2 \to \bar\theta^3 \ldots$, причому $l(\theta^1)\le l(\theta^2)\le \ldots$

Ми будуємо:

Ми знаємо:

$$\bar\theta^{j+1} = \arg\max_{\bar\theta^*} \sum_{i=1}^{n} \prod_{\bar k\in K^T} p(\bar k/\bar x^i;\theta^j)\cdot \log p(x^i;\bar k;\bar\theta^*) =$$

підставили

$$= \arg\max_{\bar\theta^*} \sum_{i=1}^{n} \sum_{\bar k\in K^T} p(\bar k/\bar x^i;\bar\theta^j)\cdot \log\left[p(\bar x^i/\bar k;\bar\theta^*)\cdot p(\bar k)\right] = \left\{\begin{array}{l}\text{візьмемо}\\ \text{логарифм}\\ \text{(розкладемо)}\end{array}\right.$$

$$= \arg\max_{\bar\theta^*} \sum_{i=1}^{n} \sum_{\bar k\in K^T} p(\bar k/\bar x^i;\theta^j)\cdot\left[\log p(x^i/\bar k;\bar\theta^*) + \log p(\bar k)\right] =$$

$=$ {після перемноження доданки без θ^* викидаємо, бо $\arg\max_{\bar\theta^*}$} $=$

$$= \arg\max_{\bar\theta^*} \sum_{i=1}^{n} \sum_{\bar k\in K^T} p(\bar k/\bar x^i;\bar\theta^j)\cdot \log \underbrace{p(\bar x^i/\bar k;\bar\theta^*)}_{\text{замінимо це через } (\blacktriangle)} =$$

$$= \arg\max_{\bar\theta^*} \sum_{i=1}^{n} \sum_{\bar k\in K^T} p(\bar k/\bar x^i;\bar\theta^j)\cdot \log \prod_{t\in T} p(x^i_t/k_t;\theta^*_{k_t}) = \left\{\begin{array}{l}\text{внесемо}\\ \log\end{array}\right\} =$$

$$= \arg\max_{\bar\theta^*} \sum_{i=1}^{n} \sum_{\bar k\in K^T} p(\bar k/\bar x^i;\bar\theta^j)\cdot \sum_{t\in T} \log p(x^i_t/k_t;\theta^*_{k_t}) =$$

взяли тільки по k_t.

$$= \arg\max_{\bar\theta^*} \sum_{t\in T}\sum_{i=1}^{n}\sum_{k\in K}\sum_{\substack{\bar k\in K^T\\ k_t=k}} p(\bar k/\bar x^i;\bar\theta^j)\cdot \log p(x^i_t/k;\theta^*_{k_t}) =$$

$$= \arg\max_{\theta^*_1,\theta^*_2,\ldots,\theta^*_{|K|}} \sum_{k\in K}\sum_{t\in T}\sum_{i=1}^{n} \sum_{\substack{\bar k\in K^T\\ k_t=k}} p(\bar k/\bar x^i;\theta^j)\cdot \log p(x^i_t/k;\theta^*_k)$$

ймовірність мітки k в об'єкті t позначимо як $p_{k_t}(k/\bar x^i;\bar\theta^j)$

$$\Rightarrow \Theta_k^{j+1} = \underset{\Theta_k^*}{\arg\max} \sum_{t \in T} \sum_{i=1}^{n} \underbrace{p_{k_t}(k / \bar{x}^i; \Theta^j)}_{\text{пошук } p \text{ це } (+;\cdot) \text{ задача}} \cdot \log p(x_t^i / k; \Theta_k^*)$$

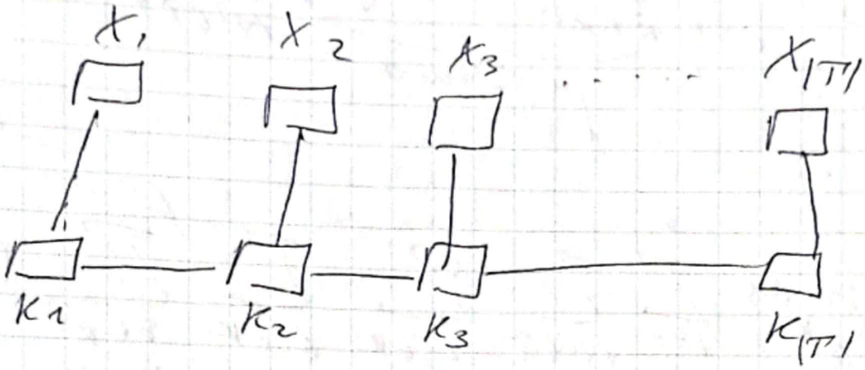
Грубе наближення

$$p_{k_t}(k / \bar{x}; \bar{\Theta}) = \begin{cases} 1, & k = k_t^* \\ 0, & \text{інакше} \end{cases}$$

$$\bar{k}^* = \underset{\bar{k}}{\arg\max}\, p(\bar{k} / \bar{x}; \Theta)$$

Лекция остання.

Есть наблюдаемые



Есть ненаблюдаемые

Известно $p(x_i/k_i)$

Неизвестно $p(\bar{k}) = \frac{1}{Z(g)} \exp\left[\sum_{tt' \in T} g_{tt'}(k_t, k_{t'})\right]$

g – неизвестны, мы бы хотели их узнать:

$$\sum_{\bar{k} \in K^T} p(\bar{k}) = 1 \Rightarrow Z(\bar{g}) = \sum_{\bar{k} \in K^T} \exp\left[\sum_{tt' \in T} g_{tt'}(k_t, k_{t'})\right]$$

$$\bar{x}^1, \bar{x}^2, \bar{x}^3, \dots, \bar{x}^n \in (X^T)^n$$

Применим EM алгоритм.

$$p(\bar{k}, \bar{x}) = \prod_{t \in T} p(x_t/k_t) \cdot \frac{1}{Z(g)} \exp\left(\sum_{tt' \in T} g_{tt'}(k_t, k_{t'})\right)$$

$$g^j \to g^{j+1}$$

тільки g^*

$$g^{j+1} = \arg\max_{g^*} \sum_{i=1}^{n} \sum_{\bar{k} \in K^T} p(\bar{k}/\bar{x}^i; \bar{g}^j) \cdot \log p(\bar{k}, \bar{x}; \bar{g}^*) =$$

$$= \arg\max_{\bar{g}^*} \sum_{i=1} \sum_{\bar{k} \in K^T} p(\bar{k}/\bar{x}^i, g^j) \Big[\underbrace{\cancel{\sum_{t \in T} \log p(x_t/k_t)}}_{\text{не зал від } g^*} +$$

$$+ \sum_{tt' \in T} g^*_{tt'}(k_t, k_{t'}) - \log Z(\bar{g}^*) \Big] =$$ } розкриємо?

$$= \arg\max_{\bar{g}^*} \sum_{i=1}^{n} \sum_{\bar{k} \in K^T} p(\bar{k}/\bar{x}^i; \bar{g}^j) \sum_{tt' \in T} g^*_{tt'}(k_t, k_{t'}) - \sum_{i=1}^{n} \sum_{\bar{k} \in K^T} p(\bar{k}/\bar{x}^i; \bar{g}^j)$$

вынесем за $\sum_{\bar{k} \in K^T}$

$$\cdot \underbrace{\log Z(\bar{g}^*)}_{\text{не зависит от } \bar{k} \text{ и } i} =$$

$$= \arg\max_{\bar g^*} \sum_{i=1}^{n} \sum_{\bar k \in K^T} p(\bar k / \bar x^i ; \bar g^j) \cdot \sum_{tt' \in T} g^*_{tt'}(k_t, k_{t'}) - n \log Z(\bar g^*) =$$

розбіймо

для k_1, k_2

$$= \arg\max_{g^*} \sum_{i=1}^{n} \sum_{tt' \in T} \sum_{k_1 \in K} \sum_{k_2 \in K} \sum_{\substack{\bar k \in K^T \\ k_t = k_1 \\ k_{t'} = k_2}} p(\bar k / \bar x^i ; \bar g^j) \, g^*_{tt'}(k_t, k_{t'}) -$$

$$- n \log Z(\bar g^*) =$$

$$= \arg\max_{g^*} \sum_{i=1}^{n} \sum_{tt' \in T} \sum_{k_1 \in K} \sum_{k_2 \in K} g^*_{tt'}(k_1, k_2) \sum_{\substack{\bar k \in K^T \\ k_t = k_1 \\ k_{t'} = k_2}} p(\bar k / \bar x^i ; \bar g^j) - n \log Z$$

обозначим

$\alpha_{tt'}(k_1, k_2 / \bar x^i ; \bar g^j)$

вер-ть, что в объекте x

метка k_t задана k_1

$k_{t'}$ задано k_2

$$= \arg\max_{g^*} \sum_{i=1}^{n} \sum_{tt' \in T} \sum_{k_1 \in K} \sum_{k_2 \in K} g^*_{tt'}(k_1, k_2) \, \alpha_{tt'}(k_1, k_2 / \bar x^i ; \bar g^j) -$$

$$- n \log Z(\bar g^*)$$

F (без аргумента).

$$\frac{\partial F}{\partial g^*_{tt'}(a, b)} = \sum_{i=1}^{n} \alpha_{tt'}(a, b / \bar x^i ; g^j) - n \frac{1}{Z(g^*)} \cdot \frac{\partial Z(\bar g^*)}{\partial g^*_{tt'}(a, b)} \;=$$

$$Z(\bar g) = \sum_{\bar k \in K^T} \exp\left[\sum_{tt' \in T} g_{tt'}(k_t, k_{t'})\right]$$

$$\frac{\partial Z}{\partial g^*_{tt'}(a,b)} = \{\text{только в } (a,b) \text{ не } \neq 0\} = \sum_{\substack{\bar k \in K^T \\ k_t = a \\ k_{t'} = b}} \exp\left[\sum_{tt' \in T} g_{tt'}(k_t, k_{t'})\right]$$

Подставим в $\frac{\partial F}{\partial g^*_{tt'}(a,b)}$.

$$= \sum_{i=1}^{n} \alpha_{tt'}(a,b / \bar x^i; \bar g^i) - n \cdot \sum_{\substack{\bar k \in K^T \\ k_t = a \\ k_{t'} = b}} \overbrace{\boxed{\frac{1}{Z(\bar g^*)} \exp\left[\sum_{tt' \in T} g_{tt'}(k_t, k_{t'})\right]}}^{p(\bar k; \bar g^*)}$$

$$= \sum_{i=1}^{n} \alpha_{tt'}(a,b / \bar x^i; \bar g^i) - n \sum_{\substack{\bar k \in K^T \\ k_t = a \\ k_{t'} = b}} p(\bar k; \bar g^*) = 0$$

$$\Rightarrow \underbrace{\sum_{\substack{\bar k \in K^T \\ k_t = a \\ k_{t'} = b}} p(\bar k; \bar g^*)}_{p_{tt'}(a, b, \bar g^*)} = \frac{1}{n} \underbrace{\sum_{i=1}^{n} \alpha_{tt'}(a b / \bar x^i; \bar g^i)}$$

вероятность что при заданных x в объекте t метка a в объект. t' метка b

вероятность того что в объекте t будет метка a а в t' метка b

если $n \to \infty$ выборка большая

$$p_{нн}(a, b, \bar{g}^*) = \frac{1}{n}\sum^{n} \alpha_{нн}(a, b / \bar{x}^i; \bar{g}^i) =$$

$$= \sum_{\bar{x} \in X^n} p(\bar{x}) \cdot \alpha_{нн}(a, b / \bar{x}; \bar{g}^i)$$

Scanned with CamScanner