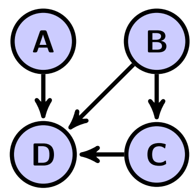
Singular Learning Theory Problems: Day 3

Shaowei Lin

23 May 2013



1. (a) Suppose we have a discrete directed graphical model whose underlying graph is the above graph. Assume that all the random variables have two states $\{0, 1\}$. Write down polynomial equations satisfied by the state probabilities $p_{ijkl} = \mathbb{P}(A = i, B = j, C = k, D = l)$ in the model which are consequences of the local Markov property.

   (b) Suppose we have a Gaussian directed graphical model with the above graph. Write down the covariance matrix Σ for the model. What are some polynomial equations satisfied by the entries $\Sigma_{ij}, i, j \in \{A, B, C, D\}$ arising from the local Markov property?

2. In the binary tree model example discussed in the lecture, we saw that the model is defined by the cumulant map $f : (\mu_a, \mu_b, \mu_c, \lambda_r, \lambda_s, \eta_s^r, \eta_c^r, \eta_a^s, \eta_b^s) \mapsto (k_a, k_b, k_c, k_{ab}, k_{ac}, k_{bc}, k_{abc})$,

$$\begin{aligned} k_a &= \mu_a, \quad k_{bc} = \tfrac{1}{4}(1 - \lambda_r^2)\eta_s^r \eta_b^s \eta_c^r, \\ k_b &= \mu_b, \quad k_{ac} = \tfrac{1}{4}(1 - \lambda_r^2)\eta_s^r \eta_a^s \eta_c^r, \\ k_c &= \mu_c, \quad k_{ab} = \tfrac{1}{4}(1 - \lambda_s^2)\eta_a^s \eta_b^s, \\ k_{abc} &= \tfrac{1}{4}(1 - \lambda_r^2)\lambda_s \eta_s^r \eta_a^s \eta_b^s \eta_c^r. \end{aligned}$$

   Find the RLCT of the map f at the point $\lambda_s = \lambda_r = 1, \mu_a = \mu_b = \mu_c = \eta_s^r = \eta_c^r = \eta_a^s = \eta_b^s = 0$.

3. Find the asymptotic coefficients (λ, θ) for the tubular volume

$$\int_{|f(\omega)| \le t} d\omega \approx C t^{\lambda} (-\log t)^{\theta - 1}$$

   where the function $f(\omega)$ is:

   (a) $f(\omega) = \omega_1^2 + \omega_2^2 + \ldots + \omega_d^2$,

   (b) $f(\omega) = \omega_1 \omega_2 \cdots \omega_d$.

4. Prove that for the restricted Boltzmann machine,

$$\mathbb{P}(h_i = 1 | v) = \mathrm{sig}(b_i + \sum_j \omega_{ij} v_j), \quad \mathrm{sig}(x) = \frac{1}{1 + e^{-x}}.$$

1 Useful formulas for computing RLCTs

Given a polynomial map $f : \Omega \to V \subset \mathbb{R}^k, f = (f_1, \ldots, f_k)$, we define the RLCT of f at an interior point $x \in \Omega \subset \mathbb{R}^d$ to be the RLCT (at the origin) of the ideal

$$I = \langle f_1(\omega + x) - f_1(x), f_2(\omega + x) - f_2(x), \ldots, f_k(\omega + x) - f_k(x)\rangle \subset \mathbb{R}[\omega_1, \ldots, \omega_d].$$

Some useful properties of RLCTs:

1. **Removing Units**.
   If $I = \langle h(\omega)g_1(\omega), g_2(\omega), \ldots, g_k(\omega)\rangle$ and $h(0) \neq 0$, then the RLCT of I is the same as the RLCT of the ideal $\langle g_1(\omega), g_2(\omega), \ldots, g_k(\omega)\rangle$ ($h(\omega)$ is removed).

2. **Monomial Ideals**.
   If I is a monomial ideal, then the RLCT is given by the Newton polyhedron method.

3. **Sum and Product Rules**.
   Given indeterminates $a_1, \ldots, a_s, b_1, \ldots, b_t$, suppose that the ideal I is generated by polynomials in $a_1, \ldots, a_s$ and the ideal J is generated by polynomials in $b_1, \ldots, b_t$. Let $\mathrm{RLCT}(I) = (\lambda_a, \theta_a)$ and $\mathrm{RLCT}(J) = (\lambda_b, \theta_b)$. Then the following formulas hold.

$$\mathrm{RLCT}(I + J) = (\lambda_a + \lambda_b,\ \theta_a + \theta_b - 1)$$

$$\mathrm{RLCT}(IJ) = \begin{cases} (\lambda_a,\ \theta_a) & \text{if } \lambda_a < \lambda_b, \\ (\lambda_b,\ \theta_b) & \text{if } \lambda_a > \lambda_b, \\ (\lambda_a,\ \theta_a + \theta_b) & \text{if } \lambda_a = \lambda_b. \end{cases}$$

4. **Global and Local RLCTs**.
   The RLCT of an ideal I globally over a set Ω is the minimum of the local RLCTs of I at each point $x \in \Omega$ (remember to translate the origin to x for each local RLCT).