A A list of symbols used

Symbols

Norms
\[ \| \|_p \] \( \ell_p \)-norm for sequences.
\[ \| \|_F \] Frobenius norm for matrices.
\[ \| \| \] Spectral norm for matrices.
\[ d_{TV}(\cdot, \cdot) \] Total variation distance between two distributions.

Sets
\( S^{n-1} \) The \( n \)-dimensional unit sphere
\( \Delta^{n-1} \) Probability simplex of dimension \( n - 1 \).
\( \Delta_{n \times (n-1)} \) Set of \( n \)-dimensional left-stochastic matrices.
\( N_0 \triangleq \{0, 1, 2, \ldots\} \) Set of nonnegative integers.
\( N_+ \triangleq \{1, 2, 3, \ldots\} \) Set of strictly positive integers.
\( \mathbb{R} \triangleq (-\infty, \infty) \) Set of reals.
\( [n] \triangleq \{1, 2, \ldots, n\} \) Set of integers 1 through \( n \).
\( \text{Perm}(n) \) All permutations of length \( n \).

Asymptotics
\( f(n) = o(g(n)) \) Small-o notation.
\( f(n) = O(g(n)) \) Big-o notation.
\( f(n) = o(g(n)) \) Little-o notation.
\( f(n) = \Omega(g(n)) \) Big-omega notation.
\( f(n) \sim g(n) \) Asymptotic equivalence.
\( f(n) = \omega(g(n)) \) Little-omega notation.
\( X_n = \Omega_p(a_n) \) Omega-p notation.
\( \forall \varepsilon: \exists \delta, N_x: \mathbb{P}[|X_n/a_n| \leq \delta] \leq \varepsilon \forall n \geq N_x \) Stochastic boundedness.
\( X_n \asymp_p (a_n) \) Equivalence-p notation.
\( \forall \varepsilon, \delta: \exists \delta_x, N_x: \mathbb{P}[|X_n/a_n| \leq \delta] \geq 1 - \varepsilon \forall n \geq N_x \) Convergence in probability.

"an absolute constant"

Block Markov Chain
\( n \in \mathbb{N}_+ \) Number of states.
\( K \in \mathbb{N}_+ \) Number of clusters.
\( V_1, \ldots, V_K \subseteq [n] \triangleq \mathcal{V} \) Clusters, and set of all states.
\( V = \bigcup_{k=1}^K V_k \), and \( V_a \cap V_b = 0 \forall a \neq b \) Relative sizes of the clusters.
\( \alpha_k \triangleq \lim_{n \to \infty} |V_k|/n \) Minimum and maximum relative sizes of the clusters.
\( \sigma: [n] \to [K] \) Cluster assignment.
\( \{X_t\}_{t \geq 0} \) Markov chain.

Transition matrices.
\( P_{x,y} \triangleq \mathbb{P}(\sigma(x) = \sigma(y) | \sigma(x) = y) \mathbb{1}[x \neq y] \forall x, y \in \mathcal{V} \) Assumed separability constant.
\( \exists \eta: \max_{a,b,c} \{p_{a,b}/p_{c,a}, p_{a,b}/p_{a,c}\} \leq \eta \) (Limiting) Equilibrium distribution.
\( \Pi_x \triangleq \lim_{t \to \infty} \mathbb{P}[X_t = x] \forall x \in \mathcal{V} \) Limiting distribution.
\( \pi_k \triangleq \lim_{n \to \infty} \sum_{x \in V_k} \Pi_x \forall k \in [K] \) Time-reversed transition matrix.
\( \delta(P) \triangleq \frac{1}{2} \sup_{x,y \in \mathcal{V}} \sum_{z \in \mathcal{V}} |P_{x,z} - P_{y,z}| \) Dobrushin’s ergodic coefficient.
0 ≤ \( t_{\text{mix}}(\varepsilon) \) ≤ \(-c_{\text{mix}} \ln \varepsilon\) Mixing time.
\( t_{\text{mix}}(\varepsilon) \triangleq \min \{ t ≥ 0 : \sup_{x \in V} d_{TV}(P_x^t, \Pi) ≤ \varepsilon \} \)
We prove the second inequality, with \( c_{\text{mix}} \) being an absolute constant.
\( \gamma_{ps} \)
Pseudo spectral gap.
\( \gamma_{ps} \triangleq \max_{i ≥ 1} \frac{1-\lambda_i(P^T)\lambda_i}{1} \)

**Information bound**
\( \mathbb{P}_P[X_0 = x_0, \ldots, X_T = x_T] \)
Probability of a sample path.
\( V^* \)
Vertex chosen uniformly at random from two clusters.
\( V^* \triangleq \text{Unif}([V_a \cup V_b], a, b \in [K], a \neq b) \)
\( \Phi \)
Probability measure of the true model.
I.e., under \( P \) and cluster assignments \( \mathcal{V}_1, \ldots, \mathcal{V}_K \).
\( \Psi \)
Probability measure of the modified model.
I.e., under \( Q \) which is a perturbation of \( P \) constructed after the random vertex \( V^* \) was moved into its own cluster.
\( Q(k, l) \)
Set of all possible change-of-measure parameters.
\( Q \triangleq \{ (q_{0, 0}, q_{0, k})_{k=0}^{K} \in (0, \infty) | q_{0, 0} = 0, \sum_{l=1}^{K} q_{0, l} = 1 \} \)
Sets of change-of-measure parameters leading to confusion between assigning to either cluster \( k, l \).
\( Q(k, l) \triangleq \{ q \in Q | I_k(q|p) = I_l(q|p) \} \neq \emptyset, k \neq l \)
\( L \triangleq \ln \frac{\mathbb{P}_Q[X_0, X_1, \ldots, X_T]}{\mathbb{P}_P[X_0, X_1, \ldots, X_T]} \)
Log-likelihood ratio.
\( I_c(q|p) \)
Leading order coefficient in an asymptotic expansion of the log-likelihood ratio.
\( I_{a,b}(q|p) \)
Deconditioned leading order coefficient in an asymptotic expanding of the log-likelihood ratio.
\( I_{a,b}(\alpha, p) \)
Separation between cluster \( a \) and \( b \).
\( J(\alpha, p) \)
Information quantities.
\( J(\alpha, p) \triangleq \min_{k \neq l} \min_{q \in Q(k, l)} I_{k,l}(q|p) \)
\( J(\alpha, p) \triangleq \min_{a \neq b} I_{a,b}(\alpha, p) \)

**Generic algorithm quantities**
\( T \in \mathbb{N}_+ \)
Observation length.
\( X_0, X_1, \ldots, X_T \)
Sample path of our Markov chain.
\( N \in \mathbb{N}_0^{n \times n} \)
Observation matrix.
Each entry contains the number of times the Markov chain jumped from \( x \) to \( y \), i.e.,
\( N_{x,y} \triangleq \sum_{t=0}^{T-1} 1[X_t = x, X_{t+1} = y] \forall x, y \in V \)
Approximated cluster assignments.
\( \gamma_{opt} \in \min_{\gamma \in \text{Perm}(K)} \bigcup_{k=1}^{K} \hat{V}_\gamma(k) \setminus \check{V}_k \)
Permutation that minimizes the overlap between the true clusters and a cluster assignment.
\( \mathcal{E} \subset \mathcal{V} \)
Set of misclassified vertices.
\( \mathcal{E} \triangleq \cup_{k=1}^{K} \hat{V}_{\gamma_{opt}(k)} \setminus \check{V}_k \)

**Regime terminologies**
“asymptotically accurate detection”
\( \mathbb{E}_P[|\mathcal{E}|] = o(n) \)
\( \mathbb{E}_P[|\mathcal{E}|] = o(1) \)
\( T = \omega(n \ln n) \)
\( T \sim cn \ln n \) for some absolute constant \( c > 0 \)
\( \omega(n) = T = o(n \ln n) \)

**Spectral clustering algorithm**
\( \Gamma \subseteq \mathcal{V} \)
Truncated vertices.
This set is obtained from \( \mathcal{V} \) by removing the \( \lfloor n \exp(-(T/n) \ln (T/n)) \rfloor \) states with the highest numbers of visits in the observed sample path of length \( T \).
\( \hat{N}_\Gamma \)
Truncated observation matrix.
This matrix is obtained from \( \hat{N} \) by setting all entries on the rows and columns corresponding to setates not in \( \Gamma \) to zero.
\( U \Sigma V^T \)
Singular value decomposition of \( \hat{N}_\Gamma \).
\( \sigma_1 ≥ \sigma_2 ≥ \cdots ≥ \sigma_n ≥ 0 \)
Singular values of \( \hat{N}_\Gamma \).
\( \hat{R} \triangleq \sum_{k=1}^{K} \sigma_k U_{., k} V_{., k}^T \)
Best rank-\( K \) approximation of \( \hat{N}_\Gamma \).
\[ \mathcal{N}_x \]

Iteratively constructed cluster centers.

\[ z_1^*, \ldots, z_K^* \in \mathcal{V} \]

*Cluster improvement algorithm*

\[ \hat{p}, \hat{\pi}, \hat{\alpha} \]

\[ u_k^{(t)}(c) \]

\[ \mathcal{H} \subseteq \mathcal{V} \]

\[ \mathcal{N}_x \triangleq \left\{ y \in \mathcal{V} \left| \sqrt{\| \hat{R}_{x,-} - \hat{R}_{y,-} \|^2 + \| \hat{R}_{x,y} - \hat{R}_{y,y} \|^2} \leq \frac{1}{\mathcal{P}} \right\} \right. \]

\[ \left( \frac{\mathcal{P}}{3} \right)^{3/2} \left( \ln \frac{\mathcal{P}}{3} \right)^{4/3} \]

Iteratively constructed cluster centers.

Approximated BMC parameters.

Approximated difference between two log-likelihood functions.

\[ u_k^{(t)}(c) \triangleq \left\{ \sum_{k=1}^{K} \left( \mathcal{N}_{x,y}^{(t)} \ln \hat{p}_{x,k} + \mathcal{N}_{y,y}^{(t)} \ln \frac{\hat{p}_{y,k}}{\hat{\pi}_k} \right) - \frac{\mathcal{P}}{3} \right\} \]

Set of well-behaved vertices.