Estimation of Markov Chain via Rank-constrained Likelihood

Xudong Li [1]   Mengdi Wang [1]   Anru Zhang [2]

Abstract

This paper studies the estimation of low-rank Markov chains from empirical trajectories. We propose a non-convex estimator based on rank-constrained likelihood maximization. Statistical upper bounds are provided for the Kullback-Leiber divergence and the ℓ_2 risk between the estimator and the true transition matrix. The estimator reveals a compressed state space of the Markov chain. We also develop a novel DC (difference of convex function) programming algorithm to tackle the rank-constrained non-smooth optimization problem. Convergence results are established. Experiments show that the proposed estimator achieves better empirical performance than other popular approaches.

1. Introduction

In scientific studies, engineering systems, and social environments, one often has to interact with complex processes with uncertainties, collect data, and learn to make predictions and decisions. A critical question is how to model the unknown random process from data. To answer this question, we particularly focus on discrete Markov chains with a finite but potentially huge state space. Suppose we are given a sequence of observations generated by an unknown Markov process. The true transition matrix is unknown, and the physical states of the system and its law of transition is hidden under massive noisy observations. We are interested in the estimation and state compression of Markov processes, i.e., to identify compact representations of the state space and recover a reduced-dimensional Markov model.

To be more specific, we focus on the estimation of a class

[1]Department of Operations Research and Financial Engineering, Princeton University, Sherrerd Hall, Princeton, NJ 08544 [2]Department of Statistics, University of Wisconsin-Madison, Madison, WI 53706. Correspondence to: Xudong Li <xudongl@princeton.edu>, Mengdi Wang <mengdiw@princeton.edu>, Anru Zhang <anruzhang@stat.wisc.edu>.

of Markov chains with low-rank transition matrices in this paper. Low-rankness is ubiquitous in practice. For example, the random walk of a taxi turns out to correspond to a nearly low-rank Markov chain. Each taxi trip can be viewed as a sample transition of a city-wide random walk that characterizes the traffic dynamics. It is possible to analyze records of taxi trips to reveal the traffic network of a metropolitan (Liu et al., 2012; Benson et al., 2017). For another example, control and reinforcement learning, which are typically modeled as Markov decision processes, suffer from the curse of dimensionality (Sutton & Barto, 1998). It is of vital interest to identify the compressed representation of the "state" of game, in order to reduce the dimensionality of policy and value functions. Multiple efforts show that as long as a reduced model or a compact state representation is given, one can solve high-dimensional control problems in sample-efficient ways; see (Singh et al., 1995; Van Roy, 2006; Malek et al., 2014) for examples.

The low-rank Markov chain is a basic model for analyzing data generated by stochastic processes. It is closely related to a variety of latent-variable models, including Markov process with rich observations (Azizzadenesheli et al., 2016), state aggregation models (Singh et al., 1995) and hidden Markov models (Hsu et al., 2012). Low-rankness of Markov chains can be used to recover network partition under some lumpability assumption (E et al., 2008). In particular, low-rank Markov chain can be viewed as a special form of hidden Markov model where the observations contain all the information of the hidden states.

In this article, we propose a rank-constrained maximum likelihood approach for transition matrix estimation from empirical trajectories. The statistical guarantees is developed for the proposed estimator. Especially, the upper bounds for the Kullback-Leibler (KL) divergence between the estimator and the true transition matrix as well as the ℓ_2 risk are established. We further provide a minimax lower bound to show that the proposed estimator is near rate-optimal within a class of low-rank Markov chains. A related recent work (Zhang & Wang, 2018) considered the low-rank estimation by a spectral method and established upper bound for the total variation. Compared to the existing results, our maximum likelihood estimator and the KL-divergence bounds appear to be the first ones of this type.

The proposed estimator requires solving an optimization

problem with both rank and polyhedral constraints. Due to the presence of non-convex rank constraint, the optimization problem is difficult – there is no efficient approach that directly applies to such a non-convex problem, to the best of our knowledge. In this paper, we use a penalty approach to relax the rank constraint and transform the original problem into a DC (difference of convex functions) programming one. Furthermore, we develop a corresponding DC algorithm to solve the problem. The algorithm proceeds by solving a sequence of inner subproblems, for which we develop an efficient subroutine based on multi-block alternating direction method of multipliers (ADMM). As a byproduct of this research, we develop a new class of DC algorithms and a unified convergence analysis for solving non-convex non-smooth problems.

Finally, the performance of the proposed estimator and algorithm is verified through simulation studies.

2. Related literatures

Model reduction for complicated systems has a long history that traces back to variable-resolution dynamic programming (Moore, 1991) and state aggregation for decision process (Sutton & Barto, 1998). In particular, (Deng et al., 2011; Deng & Huang, 2012) considered low-rank reduction of Markov models with explicitly known transition probability matrix, but not the estimation of the reduced models.

Low-rank models and spectral methods have been proved powerful in extracting features from high-dimensional data, with numerous applications including network analysis (E et al., 2008), community detection (Newman, 2013), ranking (Negahban et al., 2016), product recommendation (Keshavan et al., 2010) and many more. Estimation of hidden Markov models is related to spectral matrix/tensor decomposition, which has been considered in (Hsu et al., 2012). Later, (Anandkumar et al., 2014) studied a more general tensor-decomposition-based spectral approach, which can be applied to a variety of latent variable models including hidden Markov models. (Huang et al., 2016) considered the estimation of a rank-two probabilistic matrix from a few number of independent samples. Estimation of low-rank Markov chain was considered for the first time in (Zhang & Wang, 2018). Particularly, a spectral method via truncated singular value decomposition was introduced and the upper and lower error bounds in total variation were established. (Yang et al., 2017) developed an online stochastic gradient method for computing the leading singular space of a transition matrix from random walk data. The estimation of discrete distributions in the form of vectors is another line of related research, where the procedure and estimation risk has been considered in both classic and recent literature (Steinhaus, 1957; Lehmann & Casella, 2006; Han et al., 2015).

Form the perspective of optimization, DC programming is a backbone in handling the rank-constrained problem. First introduced by (Pham Dinh & Le Thi, 1997), the DC algorithm has become a prominent tool for handling a class of nonconvex optimization problems (see also (Pham Dinh & Le Thi, 2005; Le Thi et al., 2012; 2017; Le Thi & Pham Dinh, 2018)). In particular, (Van Dinh et al., 2015; Wen et al., 2017) considered the majorized indefinite-proximal DC algorithm, which is closely related to the optimization methods in this paper. However, both (Van Dinh et al., 2015; Wen et al., 2017) used the majorization technique with restricted choices of majorants, neither considered the introduction of the indefinite proximal terms, and (Wen et al., 2017) further requires the smooth part in the objective to be convex. A more flexible algorithmic framework and a unified convergence analysis is provided in Section 5.4.

3. Low-rank Markov chains and latent variable models

Consider a discrete-time Markov chain on p states $\{s_1, \ldots, s_p\}$ with transition matrix $\mathbf{P} \in \Re^{p \times p}$. Then $\mathbf{P}$ satisfies $P_{ij} = P(s_j | s_i)$, $0 \le P_{ij} \le 1$, and $\sum_{j=1}^{p} P_{ij} = 1$ for $1 \le i, j \le p$. Let μ be the corresponding stationary distribution. Suppose $\mathrm{rank}(\mathbf{P}) = r$. Motivated by the examples that we will discuss below, we focus on the case where $r \ll p$, i.e., the rank of $\mathbf{P}$ is significantly smaller than the total number of states.

The first example of low-rank Markov chain in reduced-order models is the state aggregation – a basic model for describing complicated systems (Bertsekas, 1995; Bertsekas & Tsitsiklis, 1995). If a process admits an inherent state aggregation structure, its states can be clustered into a small number of disjoint subsets, such that states from the same cluster have identical transition distribution. We say that a Markov chain is state aggregatable if there exists a partition $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_r$ such that

$$P(\cdot \mid s) = P(\cdot \mid s'), \qquad \forall s, s' \in \mathcal{S}_i, \ i \in [r].$$

It can be shown that the state aggregation yields a special factorization of transition matrix. Particularly, there exists $\mathbf{U}, \mathbf{V} \in \Re^{p \times r}$ such that

$$\mathbf{P} = \mathbf{U}\mathbf{V}^\top,$$

where each row of $\mathbf{U}$ has all 0's except exact one 1 and $\mathbf{V}$ is nonnegative.

Another related model is referred to as *Markov decision process with rich observations* (Azizzadenesheli et al., 2016) in the context of reinforcement learning. It is essentially a latent-variable model for Markov chains. We say that a Markov chain is a r-latent-variable model if there exists a stochastic process $\{z_t\} \subset [r]$ such that

$$P(z_t \mid s_t) = P(z_t \mid s_1, \ldots, s_t)$$

$$P(s_{t+1} \mid z_t) = P(s_{t+1} \mid s_1, \ldots, s_t, z_t).$$

The latent-variable model, another instance of low-rank Markov chains, is slightly more general than the state aggregation model. One can show that the transition matrix $\mathbf{P}$ of a r-latent-variable Markov chain has a nonnegative rank at most r and there exists $\mathbf{U}, \mathbf{V} \in \Re^{p \times r}, \tilde{\mathbf{P}} \in \Re^{r \times r}$ such that

$$\mathbf{P} = \mathbf{U}\tilde{\mathbf{P}}\mathbf{V}^\top,$$

where $\tilde{\mathbf{P}}$ is a stochastic matrix, rows of $\mathbf{U}$ and columns of $\mathbf{V}$ are vectors of probability distributions.

A third related notion is the "lumpability" of Markov chain (Buchholz, 1994). We say that a Markov chain is lumpable with respect to a partition $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_r$ if for any $i, j \in [r]$ and $s', s \in \mathcal{S}_i$,

$$P(\mathcal{S}_j \mid s) = P(\mathcal{S}_j \mid s').$$

Lumpability characterizes some partition of the state space that preserves the strong Markov property. It implies that some eigenvectors of the transition matrix is equal to the indicator functions of the subsets. However, it is a weaker assumption than state aggregation and does not necessarily imply low-rankness. As pointed out by (E et al., 2008), the lumpable partition can be recovered by clustering the singular vectors of the transition matrix.

In summary, low-rankness is a ubiquitous in a broad class of reduced-order Markov models. Estimation of low-rank transition matrices involves estimating the leading subspace, which provides a venue towards state compression.

4. Minimax rate-optimal estimation of low-rank Markov chains

Now we consider the estimation of transition matrix based on a single state-transition trajectory. Given a Markov chain $\mathbf{X} = \{X_0, X_1, \ldots, X_n\}$, we aim to estimate the transition matrix $\mathbf{P}$ given the assumption that $\text{rank}(\mathbf{P}) = r$ and $r \ll p$. For $1 \leq i, j \leq p$, denote the transition counts from state s_i to s_j by n_{ij}:

$$n_{ij} := |\{1 \leq k \leq N \mid X_{k-1} = s_i, X_k = s_j\}|.$$

Let $n_i := \sum_{j=1}^p n_{ij}$ for $i = 1, \ldots, p$ and $n = \sum_{i=1}^p n_i$.

Proposition 1. *The negative log-likelihood of $\mathbf{P}$ based on state-transition trajectory $\{x_0, \ldots, x_n\}$ is*

$$L(\mathbf{P}) := -\frac{1}{n} \sum_{i=1}^p \sum_{j=1}^p n_{ij} \log(P_{ij}). \tag{1}$$

We propose the following rank-constrained maximum likelihood estimator:

$$
\begin{aligned}
\widehat{\mathbf{P}} = \quad & \arg\min L(\mathbf{Q}) \\
s.t. \quad & \mathbf{Q}\mathbf{1}_p = \mathbf{1}_p, \ \text{rank}(\mathbf{Q}) \leq r, \\
& 0 \leq Q_{ij} \leq 1, \quad \forall\, 1 \leq i, j \leq p.
\end{aligned}
\tag{2}
$$

Then we aim to analyze the theoretical property of $\widehat{\mathbf{P}}$. Since the direct analysis of (2) for general transition matrices is technical challenging, we consider a slightly different setting, assume $\mathbf{P}$ is entry-wise bounded, and analyze the rank-constrained optimization with corresponding constraints. As we will show later in numerical analysis, such additional assumptions and constraints may not be necessary in practice.

Theorem 1 (Statistical Recovery Guarantee). *Suppose the transition matrix $\mathbf{P}$ satisfies $\text{rank}(\mathbf{P}) \leq r$, $\alpha/p \leq \mathbf{P}_{ij} \leq \beta/p$ for constants $0 < \alpha < 1 < \beta$. Consider the following programming,*

$$
\begin{aligned}
\widehat{\mathbf{P}} = \quad & \arg\min L(\mathbf{Q}) \\
s.t. \quad & \mathbf{Q}\mathbf{1}_p = \mathbf{1}_p, \ \text{rank}(\mathbf{Q}) \leq r, \\
& \alpha \leq pQ_{ij} \leq \beta \quad \forall\, 1 \leq i, j \leq p.
\end{aligned}
\tag{3}
$$

If we observe a total number of n independent state transitions generated from the stationary distribution and $n \geq Cp \log(p)$, the optimal solution $\widehat{\mathbf{P}}$ of problem (3) satisfies

$$D_{KL}(\mathbf{P}, \widehat{\mathbf{P}}) = \sum_{i,j=1}^p \mu_i P_{ij} \log(P_{ij}/\widehat{P}_{ij})$$

$$\leq C \left(\frac{pr \log p}{n} + \sqrt{\frac{\log p}{n}} \right) \tag{4}$$

$$\sum_{i=1}^p \|P_{i\cdot} - \widehat{P}_{i\cdot}\|_2^2 \leq C \left(\frac{pr \log p}{n} + \sqrt{\frac{\log p}{n}} \right) \tag{5}$$

with probability at least $1 - Cp^{-c}$. Here, $C, c > 0$ are universal constants.

Remark 1. *Note that the result in Theorem 1 applies to the data when each state transition is sampled independently from the stationary distribution of the Markov chain. In practice, samples from Markov chains are typically dependent. Then we can artificially introduce independence by sampling the dependent random walk data. Specifically, let $\tau(\epsilon)$ be the ϵ-mixing time of the Markov chain (see, e.g., (Levin & Peres, 2017)). We pick a sufficiently small ϵ and sample one transition every $\tau(\epsilon)$ transitions so that the down-sampled data are nearly independent (whose distribution is within ϵ total-variation distance from the independent data). Note that $\tau(\epsilon) \approx \tau(1/4) \log(1/\epsilon)$ scales logarithmic in $1/\epsilon$. So we can pick ϵ to be sufficiently small without deteriorating the analysis much. Therefore, the overall sample complexity in the case of random walk data is roughly $\tau(1/4)$ times the sample complexity in the independent case (up to polylog factors).*

Remark 2. *In the proof of Theorem 1, in order to bound the difference between the empirical and real loss functions for the rank-constraint estimator, we apply the concentration inequalities for empirical process and a "peeling scheme."*

In particular, we consider a partition for the set of all possible $\widehat{\mathbf{P}}$, then derive estimation loss upper bounds for each of these subsets based on concentration inequalities. The techniques used here are related to recent works on matrix completion (Negahban & Wainwright, 2012), although our problem setting, methodology, and sampling scheme are all very different from matrix completion.

Remark 3. *In Theorem 1, we use entry-wise boundedness condition α & β instead of the incoherence condition, to ensure bounded Hessian of the KL-divergence. α & β do not appear in bounds as they are treated as constants – dependence on α & β is complicated and beyond the scope of this paper. Similar condition of entry-wise boundedness was used in matrix completion (Recht, 2011, Assumption A1) and discrete distribution estimation (Kamath et al., 2015). In fact, based on (Kamath et al., 2015, Theorem 10), the entry-wise lower bound is necessary for deriving reasonable statistical recovery bounds in KL-divergence.*

Define the following set of transition matrices $\mathcal{P} = \{\mathbf{P} : \mathbf{P}1_p = 1_p, 0 \le \mathbf{P}_{ij} \le 1, \text{rank}(\mathbf{P}) \le r\}$. We further have the following lower bound results.

Theorem 2 (Lower Bound). *When $n \ge Cpr$, $\widehat{\mathbf{P}}$ is any estimator for the transition matrix $\mathbf{P}$ based on a sample trajectory of length n. Then,*

$$\inf_{\widehat{\mathbf{P}}} \sup_{\mathbf{P} \in \mathcal{P}} \mathbb{E}D_{KL}(\mathbf{P}, \widehat{\mathbf{P}}) \ge c\frac{pr}{n};$$

$$\inf_{\widehat{\mathbf{P}}} \sup_{\mathbf{P} \in \mathcal{P}} \mathbb{E} \sum_{i=1}^{p} \|\widehat{P}_{i\cdot} - P_{i\cdot}\|_2^2 \ge c\frac{pr}{n},$$

where C, c are universal constants.

Theorems 1 and 2 together show that the proposed estimator is rate-optimal up to a logarithm factor, as long as $n = O(pr \log p)$.

Remark 4. *Especially when $r = 1$, $\mathbf{P}$ can be written as $1v^\top$ for some vector $v \in \Re^p$, and then estimating $\mathbf{P}$ essentially reduces to estimating a discrete distribution from multinomial count data. Our upper bound in Theorem 1 nearly matches (up to a log factor) the classical results of discrete distribution estimation ℓ_2 risks (see, e.g. (Lehmann & Casella, 2006, Pg. 349)).*

In addition to the full transition matrix $\mathbf{P}$, the leading left and right singular subspaces of $\mathbf{P}$, say $\mathbf{U}, \mathbf{V} \in \Re^{p \times r}$, also play important roles in Markov chain analysis. For example, by performing k-means on the reliable estimations of $\mathbf{U}$ or $\mathbf{V}$ for a state aggregatable Markov chains, one can achieve good performance of state aggregation (Zhang & Wang, 2018). Based on previous discussions, one can further establish the error bounds on singular subspace estimation for Markov transition matrix.

Theorem 3. *Under the setting of Theorem 1, suppose the left and right singular subspaces of $\widehat{\mathbf{P}}$ are $\widehat{\mathbf{U}} \in \Re^{p \times r}$ and $\widehat{\mathbf{V}} \in \Re^{p \times r}$, where $\widehat{\mathbf{P}}$ is the optimizer of (3), one has*

$$\max\{\|\sin\Theta(\widehat{\mathbf{U}}, \mathbf{U})\|_F^2, \|\sin\Theta(\widehat{\mathbf{V}}, \mathbf{V})\|_F^2\}$$
$$\le \min\left\{\frac{C(pr\log p)/n + C\sqrt{(\log p)/n}}{\sigma_r^2(\mathbf{P})}, r\right\}$$

with probability at least $1 - Cp^{-c}$. Here, $\sigma_r(\mathbf{P})$ is the r-th singular value of $\mathbf{P}$ and $\|\sin\Theta(\widehat{\mathbf{U}}, \mathbf{U})\|_F = (r - \|\widehat{\mathbf{U}}^\top \mathbf{U}\|_F^2)^{1/2}$.

Remark 5. *Similarly as Theorem 2, one can develop the corresponding lower bound to show the optimality for the upper bound in Theorem 3.*

5. Optimization methods for the rank-constrained likelihood problem

In this section, we develop the optimization method for computing the rank-constrained likelihood maximizer (2). In Section 5.1, a penalty approach is applied to transform the original intractable rank-constrained problem into a DC programming problem. Then we solve this problem by a proximal DC (PDC) algorithm in Section 5.2. We discuss the solver for the subproblems involved in the proximal DC algorithm in Section 5.3. Lastly, a unified convergence analysis of a class of majorized indefinite-proximal DC (Majorized iPDC) algorithms is provided in Section 5.4.

5.1. A penalty approach for problem (2)

Recall (2) is intractable due to the non-convex rank constraint, we introduce a penalty approach to relax such a constraint, and particularly study the following general optimization problem:

$$\min\{f(\mathbf{X}) \mid \mathcal{A}(\mathbf{X}) = b, \text{rank}(\mathbf{X}) \le r\}, \quad (6)$$

where $f : \Re^{p \times p} \to (-\infty, +\infty]$ is a closed, convex, but possibly non-smooth function, $\mathcal{A} : \Re^{p \times p} \to \Re^m$ is a linear map, $b \in \Re^m$ and $r > 0$ are given data. Especially when $f(\mathbf{X}) = -\frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{p} n_{ij}\log(X_{ij}) + \delta(\mathbf{X} \mid \mathbf{X} \ge 0)$, $\mathcal{A}(\mathbf{X}) = \mathbf{X}1_p, b = 1_p$, and $\delta(\cdot \mid \mathbf{X} \ge 0)$ is the indicator function of the closed convex set $\{\mathbf{X} \in \Re^{p \times p} \mid X \ge 0\}$, the general problem (6) becomes the original rank-constraint maximum likelihood problem (2).

Given $\mathbf{X} \in \Re^{p \times p}$, let $\sigma_1(\mathbf{X}) \ge \cdots \ge \sigma_p(\mathbf{X}) \ge 0$ be the singular values of $\mathbf{X}$. Since $\text{rank}(\mathbf{X}) \le r$ if and only if $\sigma_{r+1}(\mathbf{X}) + \ldots + \sigma_p(\mathbf{X}) = \|\mathbf{X}\|_* - \|\mathbf{X}\|_{(r)} = 0$, where $\|\mathbf{X}\|_{(r)} = \sum_{i=1}^{r} \sigma_i(\mathbf{X})$ is the Ky Fan r-norm of $\mathbf{X}$, (6) can be equivalently formulated as

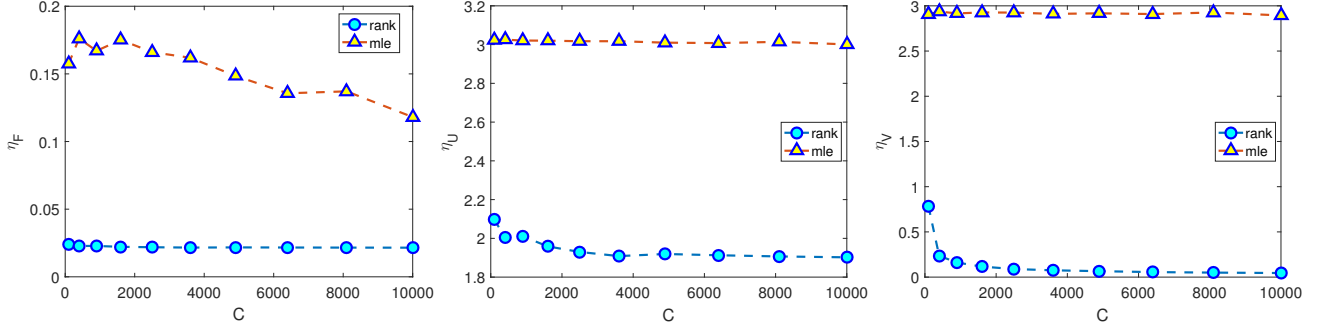$$\min\{f(\mathbf{X}) \mid \|\mathbf{X}\|_* - \|\mathbf{X}\|_{(r)} = 0, \mathcal{A}(\mathbf{X}) = b\}.$$

Figure 1. Comparison between **rank** and **mle** with accuracy measures $(\eta_F, \eta_U, \eta V)$ VS. number of state jumps $n = Crq \log(q)$.

See also (Gao & Sun, 2010, equation (29)). The penalized formulation of problem (6) is

$$\min_{\mathbf{X} \in \Re^{p \times p}} \left\{ f(\mathbf{X}) + c(\|\mathbf{X}\|_* - \|\mathbf{X}\|_{(r)}) \mid \mathcal{A}(\mathbf{X}) = b \right\}, \quad (7)$$

where $c > 0$ is a penalty parameter. Since $\| \cdot \|_{(r)}$ is convex, it is clear that the objective in problem (7) is a difference of two convex functions: $f(\mathbf{X}) + c\|\mathbf{X}\|_*$ and $c\|\mathbf{X}\|_{(r)}$, i.e., (7) is a DC program.

Let $\mathbf{X}_c^*$ be an optimal solution to the penalized problem (7). The following proposition shows that $\mathbf{X}_c^*$ is also the optimizer to (6) when it is low-rank.

Proposition 2. *If* $\mathrm{rank}(\mathbf{X}_c^*) \leq r$, *then* $\mathbf{X}_c^*$ *is also an optimal to the original problem* (6).

In practice, one can gradually increase the penalty parameter c to obtain a sufficient low rank solution $\mathbf{X}_c^*$. In our numerical experiments, we can obtain solutions with the desired rank with a properly chosen parameter c.

5.2. A PDC algorithm for penalized problem (7)

The central idea of the DC algorithm (Pham Dinh & Le Thi, 1997) is as follows: at each iteration, one approximates the concave part of the objective function by its affine majorant, then solves the resulting convex optimization problem. In this subsection, we present a variant of the classic DC algorithm for solving (7). For the execution of the algorithm, we recall that the sub-gradient of Ky Fan r-norm at a point $\mathbf{X} \in \Re^{p \times p}$ (Watson, 1993) is

$$\partial \|\mathbf{X}\|_{(r)} = \left\{ \mathbf{U} \, \mathrm{Diag}(q^*) \mathbf{V}^T \mid q^* \in \Delta \right\},$$

where $\mathbf{U}$ and $\mathbf{V}$ are the singular vectors of $\mathbf{X}$, Δ is the optimal solution set of the following problem

$$\max_{q \in \Re^p} \left\{ \sum_{i=1}^p \sigma_i(\mathbf{X}) q_i \mid \langle \mathbf{1}_p, q \rangle = r, \, 0 \leq q \leq 1 \right\}.$$

Note that one can efficiently obtain a component of $\partial \|\mathbf{X}\|_{(r)}$ by computing the SVD and picking up the SVD vectors

corresponding to the r largest singular values. After these preparations, we are ready to state the PDC algorithm for problem (7). Different from the classic DC algorithm, an additional proximal term is added to ensure the existence of solutions of subproblems (8) and the convergence of the difference of two consecutive iterations generated by the algorithm. See Theorem 4 and Remark 6 for more details.

Algorithm 1 A PDC algorithm for solving problem (7)

Given $c > 0$, $\alpha \geq 0$, and the stopping tolerance η, choose initial point $\mathbf{X}^0 \in \Re^{p \times p}$. Iterate the following steps for $k = 0, 1, \dots$:
1. Choose $\mathbf{W}_k \in \partial \|\mathbf{X}^k\|_{(r)}$. Compute

$$\mathbf{X}^{k+1} = \arg\min \left\{ \begin{array}{l} f(\mathbf{X}) + c(\|\mathbf{X}\|_* - \langle \mathbf{W}_k, \mathbf{X} - \mathbf{X}^k \rangle \\ - \|\mathbf{X}^k\|_{(r)}) + \dfrac{\alpha}{2} \|\mathbf{X} - \mathbf{X}^k\|_F^2 \\ \mid \mathcal{A}(\mathbf{X}) = b \end{array} \right\}.$$
$$(8)$$

2. If $\|\mathbf{X}^{k+1} - \mathbf{X}^k\|_F \leq \eta$, stop.

We say that $\mathbf{X}$ is a critical point of problem (7) if

$$\partial(f(\mathbf{X}) + c\|\mathbf{X}\|_*) \cap (c\partial\|\mathbf{X}\|_{(r)}) \neq \emptyset.$$

We state the following convergence results for Algorithm 1.

Theorem 4. *Let* $\{\mathbf{X}^k\}$ *be the sequence generated by Algorithm 1 and* $\alpha \geq 0$. *Then* $\{f(\mathbf{X}^k) + c(\|\mathbf{X}^k\|_* - \|\mathbf{X}^k\|_{(r)})\}$ *is a non-increasing sequence. If* $\mathbf{X}^{k+1} = \mathbf{X}^k$ *for some integer* $k \geq 0$, *then* $\mathbf{X}^k$ *is a critical point of* (7). *Otherwise, it holds that*

$$\begin{aligned} & \left(f(\mathbf{X}^{k+1}) + c(\|\mathbf{X}^{k+1}\|_* - \|\mathbf{X}^{k+1}\|_{(r)}) \right) \\ & - \left(f(\mathbf{X}^k) + c(\|\mathbf{X}^k\|_* - \|\mathbf{X}^k\|_{(r)}) \right) \\ & \leq -\frac{\alpha}{2} \|\mathbf{X}^{k+1} - \mathbf{X}^k\|_F^2. \end{aligned}$$

Moreover, any accumulation point of the bounded sequence $\{\mathbf{X}^k\}$ *is a critical point of problem* (7). *In addition, if* $\alpha > 0$, *it holds that* $\lim_{k \to \infty} \|\mathbf{X}^{k+1} - \mathbf{X}^k\|_F = 0$.
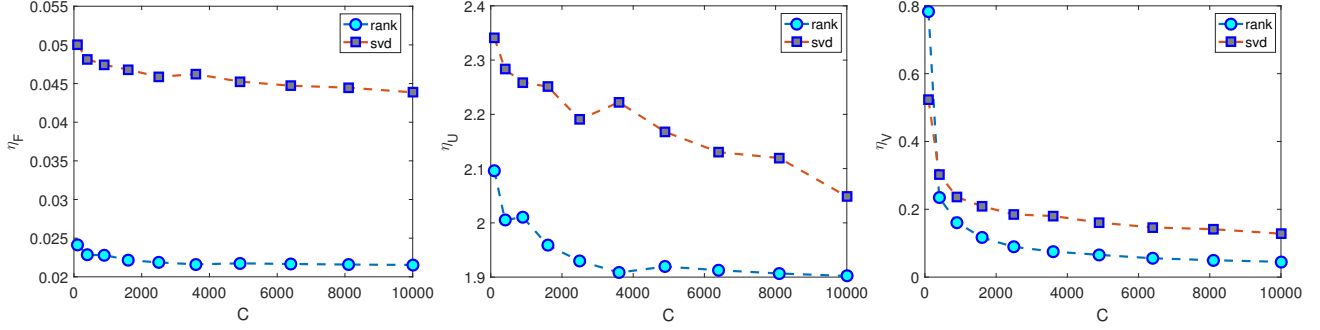
Figure 2. Comparison between **rank** and **svd** with accuracy measures $(\eta_F, \eta_U, \eta V)$ VS. number of state jumps $n = Crq\log(q)$.

Remark 6 (Selection of Parameters). *In practice, a small $\alpha > 0$ is suggested to ensure strict decrease of the objective value and convergence of $\{\|\mathbf{X}^{k+1} - \mathbf{X}^k\|_F\}$; if f is strongly convex, one achieves these nice properties even if $\alpha = 0$ based on the results of Theorem 6. The penalty parameter c can be adaptively adjusted according to the rank of the sequence generated by Algorithm 1.*

5.3. A multi-block ADMM for subproblem (8)

It is important and non-trivial to solve the convex subproblem (8) in Algorithm 1. Note that (8) is a nuclear norm penalized convex optimization problem, we propose to apply an efficient multi-block alternating direction method of multipliers (ADMM) for solving the dual of (8). A comprehensive numerical study has been conducted in (Li et al., 2016) and justifies our procedure.

Instead of working directly on (8), we study a slightly general model as follows. Given $\mathbf{W} \in \Re^{p \times p}$ and $\alpha \geq 0$, consider

$$(\mathbf{P}) \quad \min \quad f(\mathbf{X}) + \langle \mathbf{W}, \mathbf{X} \rangle + c\|\mathbf{X}\|_* + \frac{\alpha}{2}\|\mathbf{X}\|_F^2$$
$$\text{s.t.} \quad \mathcal{A}(\mathbf{X}) = b.$$

Its dual problem can be written as

$$(\mathbf{D}) \quad \min \quad f^*(-\boldsymbol{\Xi}) - \langle b, y \rangle + \frac{\alpha}{2}\|\mathbf{Z}\|_F^2$$
$$\text{s.t.} \quad \boldsymbol{\Xi} + \mathcal{A}^*(y) + \mathbf{S} + \alpha\mathbf{Z} = \mathbf{W}, \quad \|\mathbf{S}\|_2 \leq c,$$

where f^* is the conjugate function of f, $\|\cdot\|_2$ denotes the spectral norm. When we set f to be the negative likelihood function (1), f^* becomes

$$f^*(\boldsymbol{\Xi}) = \sum_{(i,j) \in \Omega} \frac{n_{ij}}{n}\left(\log\frac{n_{ij}}{n} - 1 - \log(-\Xi_{ij})\right)$$
$$+ \sum_{(i,j) \in \overline{\Omega}} \delta(\Xi_{ij} \mid \Xi_{ij} \leq 0) \quad \forall \boldsymbol{\Xi} \in \Re^{p \times p},$$

where $\Omega = \{(i,j) \mid n_{ij} \neq 0\}$ and $\overline{\Omega} = \{(i,j) \mid n_{ij} = 0\}$. Given $\sigma > 0$, the augmented Lagrangian function $\mathcal{L}_\sigma$

associated with $(\mathbf{D})$ is

$$\mathcal{L}_\sigma(\boldsymbol{\Xi}, y, \mathbf{S}, \mathbf{Z}; \mathbf{X}) = f^*(-\boldsymbol{\Xi}) - \langle b, y \rangle + \frac{\alpha}{2}\|\mathbf{Z}\|_F^2$$
$$+ \frac{\sigma}{2}\|\boldsymbol{\Xi} + \mathcal{A}^*(y) + \mathbf{S} + \alpha\mathbf{Z} - \mathbf{W} + \mathbf{X}/\sigma\|^2 - \frac{1}{2\sigma}\|\mathbf{X}\|^2.$$

We consider ADMM type methods for solving problem $(\mathbf{D})$. Since there are four separable blocks in $(\mathbf{D})$ (namely $\boldsymbol{\Xi}$, y, $\mathbf{S}$, and $\mathbf{Z}$), the direct extended ADMM is not applicable. Fortunately, the functions corresponding to blocks y and $\mathbf{Z}$ in the objective of $(\mathbf{D})$ are linear-quadratic. Thus we can apply the multi-block symmetric Gauss-Sediel based ADMM (sGS-ADMM) (Li et al., 2016). In literature (Chen et al., 2017; Ferreira et al., 2017; Lam et al., 2017; Li et al., 2016; Wang & Zou, 2018), extensive numerical experiments demonstrate that sGS-ADMM is not only convergent but also faster than the directly extended multi-block ADMM and its many other variants. Specifically, the algorithmic framework of sGS-ADMM for solving $(\mathbf{D})$ is presented in Algorithm 2. Note that when $\alpha = 0$, the computation steps corresponding to block $\mathbf{Z}$ will not be performed.

Algorithm 2 An sGS-ADMM for solving $(\mathbf{D})$

Input: initial point $(\boldsymbol{\Xi}^0, y^0, \mathbf{S}^0, \mathbf{Z}^0, \mathbf{X}^0)$, penalty parameter $\sigma > 0$, maximum iteration number K, and the step-length $\gamma \in (0, (1 + \sqrt{5})/2)$

for $k = 0$ **to** K **do**
$\quad y^{k+\frac{1}{2}} = \arg\min_y \mathcal{L}_\sigma(\boldsymbol{\Xi}^k, y, \mathbf{S}^k, \mathbf{Z}^k; \mathbf{X}^k)$
$\quad \boldsymbol{\Xi}^{k+1} = \arg\min_{\boldsymbol{\Xi}} \mathcal{L}_\sigma(\boldsymbol{\Xi}, y^{k+\frac{1}{2}}, \mathbf{S}^k, \mathbf{Z}^k; \mathbf{X}^k)$
$\quad y^{k+1} = \arg\min_y \mathcal{L}_\sigma(\boldsymbol{\Xi}^{k+1}, y, \mathbf{S}^k, \mathbf{Z}^k; \mathbf{X}^k)$
$\quad \mathbf{Z}^{k+\frac{1}{2}} = \arg\min_{\mathbf{Z}} \mathcal{L}_\sigma(\boldsymbol{\Xi}^{k+1}, y^{k+1}, \mathbf{S}^k, \mathbf{Z}; \mathbf{X}^k)$
$\quad \mathbf{S}^{k+1} = \arg\min_{\mathbf{S}} \mathcal{L}_\sigma(\boldsymbol{\Xi}^{k+1}, y^{k+1}, \mathbf{S}, \mathbf{Z}^{k+\frac{1}{2}}; \mathbf{X}^k)$
$\quad \mathbf{Z}^{k+1} = \arg\min_{\mathbf{Z}} \mathcal{L}_\sigma(\boldsymbol{\Xi}^{k+1}, y^{k+1}, \mathbf{S}^{k+1}, \mathbf{Z}; \mathbf{X}^k)$
$\quad \mathbf{X}^{k+1} = \mathbf{X}^k + \gamma\sigma(\boldsymbol{\Xi}^{k+1} + \mathcal{A}^*(y^{k+1}) + \mathbf{S}^{k+1} + \alpha\mathbf{Z}^{k+1} - \mathbf{W})$
end for

When implementing Algorithm 2, only partial SVD, which is much cheaper than full SVD, is needed as $r \ll p$. The following convergence results can be directly obtained from

(Li et al., 2016).

Theorem 5. *Suppose that the solution sets of* (**P**) *and* (**D**) *are nonempty. Let* $\{(\mathbf{\Xi}^k, y^k, \mathbf{S}^k, \mathbf{Z}^k, \mathbf{X}^k)\}$ *be the sequence generated by Algorithm 2. If* $\tau \in (0, (1 + \sqrt{5})/2)$, *then the sequence* $\{(\mathbf{\Xi}^k, y^k, \mathbf{S}^k, \mathbf{Z}^k)\}$ *converges to an optimal solution of* (**D**) *and* $\{\mathbf{X}^k\}$ *converges to an optimal solution of* (**P**).

5.4. A unified analysis for the majorized iPDC algorithm

Due to the presence of the proximal term $\frac{\alpha}{2}\|\mathbf{X} - \mathbf{X}^k\|^2$ in Algorithm 1, the classical DC analyses cannot be applied directly. Hence, in this subsection, we provide a unified convergence analysis for the majorized indefinite-proximal DC (majorized iPDC) algorithm which includes Algorithm 1 as a special instance. Let $\mathbb{X}$ be a finite-dimensional real Euclidean space endowed with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\|\cdot\|$. Consider the following optimization problem

$$\min_{x \in \mathbb{X}} \ \theta(x) \triangleq g(x) + p(x) - q(x), \qquad (9)$$

where $g : \mathbb{X} \to \Re$ is a continuously differentiable function (not necessarily convex) with a Lipschitz continuous gradient and Lipschitz modulus $L_g > 0$, i.e.,

$$\|\nabla f(x) - \nabla f(x')\| \le L_g \|x - x'\| \quad \forall x, x' \in \mathbb{X},$$

$p : \mathbb{X} \to (-\infty, +\infty]$ and $q : \mathbb{X} \to (-\infty, +\infty]$ are two proper closed convex functions. It is not difficult to observe that penalized problem (7) is a special instance of problem (9). For general model (9), one can only expect the DC algorithm converges to a critical point $\bar{x} \in \mathbb{X}$ of (9) satisfying

$$(\nabla g(\bar{x}) + \partial p(\bar{x})) \cap \partial q(\bar{x}) \ne \emptyset.$$

Since g is continuously differentiable with Lipschitz continuous gradient, there exists a self-adjoint positive semidefinite linear operator $\mathcal{G} : \mathbb{X} \to \mathbb{X}$ such that for any $x, x' \in \mathbb{X}$,

$$g(x) \le \widehat{g}(x; x') \triangleq g(x') + \langle \nabla g(x'), x - x' \rangle + \frac{1}{2}\|x - x'\|_{\mathcal{G}}^2.$$

The majorized iPDC algorithm for solving (9) is presented in Algorithm 3. We further provide the following convergence results.

Theorem 6. *Assume that* $\inf \theta > -\infty$. *Let* $\{x^k\}$ *be the sequence generated by Algorithm 3. If* $x^{k+1} = x^k$ *for some* $k \ge 0$, *then* x^k *is a critical point of* (9). *If* $\mathcal{G} + 2\mathcal{T} \succeq 0$, *then any accumulation point of* $\{x^k\}$, *if exists, is a critical point of* (9). *In addition, if* $\mathcal{G} + 2\mathcal{T} \succ 0$, *it holds that* $\lim_{i \to \infty} \|x^{k+1} - x^k\| = 0$.

The proof of Theorem 6 and more discussions are provided in the supplementary material.

Algorithm 3 A majorized indefinite-proximal DC algorithm for solving problem (9)

Given initial point $x^0 \in \mathbb{X}$ and stopping tolerance η, choose a self-adjoint linear operator $\mathcal{T} : \mathbb{X} \to \mathbb{X}$ and $\mathcal{G}$. Iterate the following steps for $k = 0, 1, \dots$:
1. Choose $\xi^k \in \partial q(x^k)$. Compute

$$x^{k+1} \in \arg\min_{x \in \mathbb{X}} \left\{ \widehat{\theta}(x; x^k) + \frac{1}{2}\|x - x^k\|_{\mathcal{T}}^2 \right\}, \quad (10)$$

where $\widehat{\theta}(x; x^k) \triangleq \widehat{g}(x; x^k) + p(x) - (q(x^k) + \langle x - x^k, \xi^k \rangle)$.
2. If $\|x^{k+1} - x^k\| \le \eta$, stop.

---

6. Simulation Results

We first generate the rank-r transition matrix as $\mathbf{P} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, where $\mathbf{U}, \mathbf{V} \in \Re^{p \times r}$ have orthonormal columns and the diagonal matrix $\mathbf{\Sigma} = \mathrm{diag}(\sigma_i) \in \Re^{r \times r}$ with $\sigma_i > 0$ for $i = 1, \dots, r$. Then we simulate a Markov chain trajectory of length $n = Crp \log(p)$. Here, $r = 10$, $p = 500$ and C is a varying constant from 10^2 to 10^4.

We compare four estimation procedures, i.e., maximum likelihood estimator (**mle**), truncated SVD estimator (**svd**) (Zhang & Wang, 2018), nuclear norm penalized estimator (**nu**), and our rank-constrained maximum likelihood estimator (**rank**). For each estimator $\widehat{\mathbf{P}}$, let $\widehat{\mathbf{U}}$ and $\widehat{\mathbf{V}}$ be the leading r left and right singular vectors of $\widehat{\mathbf{P}}$, respectively. We measure the performance of $\widehat{\mathbf{P}}$ through three quantities:

$$\eta_F = \|\mathbf{P} - \widehat{\mathbf{P}}\|_F / \sqrt{p}, \quad \eta_U = \|\sin \Theta(\mathbf{U}, \widehat{\mathbf{U}})\|_F,$$
$$\eta_V = \|\sin \Theta(\mathbf{V}, \widehat{\mathbf{V}})\|_F.$$

The comparison between **mle** and **rank** is presented in Figure 1. As one can observe, the empirical error of our rank-constrained maximum likelihood estimator is significantly smaller than the plain maximum likelihood estimator.

From Figure 2, one can see that our rank-constrained estimator outperforms **svd** in terms of all three accuracy measurements. Recently, (Zhang & Wang, 2018) studied the **svd** approach and developed the total variation error bounds. The rank-constrained estimator and the KL-divergence error bounds studied in this paper are and harder to analyze and more meaningful for discrete distribution estimation.

In the implementation of nuclear norm penalized estimator **nu**, the penalty parameter is chosen via 5-fold cross-validation. The comparison between **rank** and **nu** is plotted in Figure 3. Clearly, the estimation error of **rank** is much smaller than that of **nu**. In fact, as one can see from Algorithm 1, the nuclear norm penalized estimator is in fact a single step in the procedure of our rank-constrained estimator, and **rank** gradually refines based on **nu**. Theoretically speaking, one could expect that, similar to the matrix re-
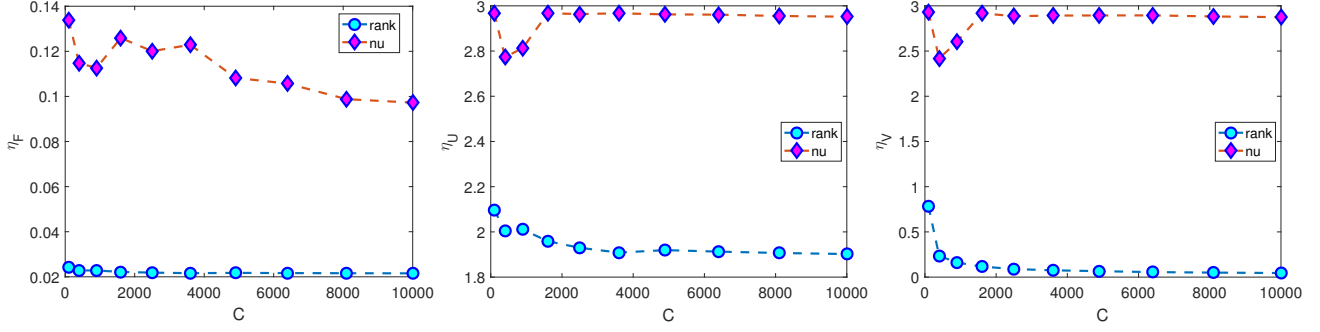
Figure 3. Comparison between **rank** and **nu** with accuracy measures $(\eta_F, \eta_U, \eta V)$ VS. number of state jumps $n = Crq\log(q)$.

covery (Recht et al., 2010), the rank-constrained estimator enjoys better statistical performances and needs weaker assumption over the nuclear norm penalized estimator. Moreover, Theorems 1 and 2 indicate that the rank-constrained maximum likelihood estimator is statistically optimal for the estimation of Markov Chains. However, such a result is not clear for the nuclear norm approach.

7. Conclusion

This paper studies the recovery and state compression of low-rank Markov chains from empirical trajectories via a rank-constrained likelihood approach. We provide statistical upper bounds for the ℓ_2 risk and Kullback-Leibler divergence between the estimator and the true probability transition matrix for the proposed estimator. Then, a novel DC programming algorithm is developed to solve the associated rank-constrained optimization problem. The proposed algorithm non-trivially combines several recent optimization techniques, such as the penalty approach, the proximal DC algorithm, and the multi-block sGS-ADMM. We further study a new class of majorized indefinite-proximal DC algorithms for solving general non-convex non-smooth DC programming problems and provide a unified convergence analysis. Experiments on simulated data illustrate the merits of our approach.

References

Anandkumar, Animashree, Ge, Rong, Hsu, Daniel, Kakade, Sham M, and Telgarsky, Matus. Tensor decompositions for learning latent variable models. *The Journal of Machine Learning Research*, 15(1):2773–2832, 2014.

Azizzadenesheli, Kamyar, Lazaric, Alessandro, and Anandkumar, Animashree. Reinforcement learning in rich-observation mdps using spectral methods. *arXiv preprint arXiv:1611.03907*, 2016.

Benson, Austin R, Gleich, David F, and Lim, Lek-Heng. The spacey random walk: A stochastic process for higher-order data. *SIAM Review*, 59(2):321–345, 2017.

Bertsekas, Dimitri P. *Dynamic Programming and Optimal Control*, volume 1. Athena scientific Belmont, MA, 1995.

Bertsekas, Dimitri P and Tsitsiklis, John N. Neuro-dynamic programming: an overview. In *Decision and Control, 1995., Proceedings of the 34th IEEE Conference on*, volume 1, pp. 560–564. IEEE, 1995.

Buchholz, Peter. Exact and ordinary lumpability in finite Markov chains. *Journal of applied probability*, 31(01): 59–75, 1994.

Chen, Liang, Sun, Defeng, and Toh, Kim-Chuan. An efficient inexact symmetric Gauss–Seidel based majorized ADMM for high-dimensional convex composite conic programming. *Mathematical Programming*, 161(1-2): 237–270, 2017.

Deng, Kun and Huang, Dayu. Model reduction of Markov chains via low-rank approximation. In *American Control Conference (ACC), 2012*, pp. 2651–2656. IEEE, 2012.

Deng, Kun, Mehta, Prashant G, and Meyn, Sean P. Optimal Kullback-Leibler aggregation via spectral theory of Markov chains. *IEEE Transactions on Automatic Control*, 56(12):2793–2808, 2011.

E, Weinan, Li, Tiejun, and Vanden-Eijnden, Eric. Optimal partition and effective dynamics of complex networks. *Proceedings of the National Academy of Sciences*, 105 (23):7907–7912, 2008.

Ferreira, José FS Bravo, Khoo, Yuehaw, and Singer, Amit. Semidefinite programming approach for the quadratic assignment problem with a sparse graph. *Computational Optimization and Applications*, pp. 1–36, 2017.

Gao, Yan and Sun, Defeng. A majorized penalty approach for calibrating rank constrained correlation matrix problems. *technical reprot*, 2010.

Han, Yanjun, Jiao, Jiantao, and Weissman, Tsachy. Minimax estimation of discrete distributions under ℓ_1 loss. *IEEE Transactions on Information Theory*, 61(11):6343–6354, 2015.

Hsu, Daniel, Kakade, Sham M, and Zhang, Tong. A spectral algorithm for learning hidden Markov models. *Journal of Computer and System Sciences*, 78(5):1460–1480, 2012.

Huang, Qingqing, Kakade, Sham M, Kong, Weihao, and Valiant, Gregory. Recovering structured probability matrices. *arXiv preprint arXiv:1602.06586*, 2016.

Kamath, Sudeep, Orlitsky, Alon, Pichapati, Dheeraj, and Suresh, Ananda Theertha. On learning distributions from their samples. In *Conference on Learning Theory*, pp. 1066–1100, 2015.

Keshavan, Raghunandan H, Montanari, Andrea, and Oh, Sewoong. Matrix completion from a few entries. *IEEE Transactions on Information Theory*, 56(6):2980–2998, 2010.

Lam, Xin Yee, Marron, JS, Sun, Defeng, and Toh, Kim-Chuan. Fast algorithms for large scale generalized distance weighted discrimination. *Journal of Computational and Graphical Statistics*, (just-accepted), 2017.

Le Thi, Hoai An and Pham Dinh, Tao. DC programming and DCA: thirty years of developments. *Mathematical Programming*, pp. 1–64, 2018. doi: 10.1007/s10107-018-1235-y.

Le Thi, Hoai An, Pham Dinh, Tao, and Van Ngai, Huynh. Exact penalty and error bounds in DC programming. *Journal of Global Optimization*, 52(3):509–535, 2012.

Le Thi, Hoai An, Le, Hoai Minh, Phan, Duy Nhat, and Tran, Bach. Stochastic DCA for the large-sum of non-convex functions problem and its application to group variable selection in classification. In *International Conference on Machine Learning*, pp. 3394–3403, 2017.

Lehmann, Erich L and Casella, George. *Theory of point estimation*. Springer Science & Business Media, 2006.

Levin, David A and Peres, Yuval. *Markov chains and mixing times*, volume 107. American Mathematical Soc., 2017.

Li, Xudong, Sun, Defeng, and Toh, Kim-Chuan. A Schur complement based semi-proximal ADMM for convex quadratic conic programming and extensions. *Mathematical Programming*, 155(1-2):333–373, 2016.

Liu, Yu, Kang, Chaogui, Gao, Song, Xiao, Yu, and Tian, Yuan. Understanding intra-urban trip patterns from taxi trajectory data. *Journal of geographical systems*, 14(4):463–483, 2012.

Malek, Alan, Abbasi-Yadkori, Yasin, and Bartlett, Peter. Linear programming for large-scale Markov decision problems. In *International Conference on Machine Learning*, pp. 496–504, 2014.

Moore, Andrew W. Variable resolution dynamic programming: Efficiently learning action maps in multivariate real-valued state-spaces. In *Machine Learning Proceedings 1991*, pp. 333–337. Elsevier, 1991.

Negahban, Sahand and Wainwright, Martin J. Restricted strong convexity and weighted matrix completion: Optimal bounds with noise. *Journal of Machine Learning Research*, 13(May):1665–1697, 2012.

Negahban, Sahand, Oh, Sewoong, and Shah, Devavrat. Rank centrality: Ranking from pairwise comparisons. *Operations Research*, 65(1):266–287, 2016.

Newman, Mark EJ. Spectral methods for community detection and graph partitioning. *Physical Review E*, 88(4):042822, 2013.

Pham Dinh, Tao and Le Thi, Ho ai. Convex analysis approach to DC programming: Theory, algorithms and applications. *Acta Mathematica Vietnamica*, 22(1):289–355, 1997.

Pham Dinh, Tao and Le Thi, Ho ai. The DC (difference of convex functions) programming and DCA revisited with DC models of real world nonconvex optimization problems. *Annals of operations research*, 133(1-4):23–46, 2005.

Recht, Benjamin. A simpler approach to matrix completion. *Journal of Machine Learning Research*, 12(Dec):3413–3430, 2011.

Recht, Benjamin, Fazel, Maryam, and Parrilo, Pablo A. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM review*, 52(3):471–501, 2010.

Singh, Satinder P, Jaakkola, Tommi, and Jordan, Michael I. Reinforcement learning with soft state aggregation. In *Advances in neural information processing systems*, pp. 361–368, 1995.

Steinhaus, Ho. The problem of estimation. *The Annals of Mathematical Statistics*, 28(3):633–648, 1957.

Sutton, Richard S and Barto, Andrew G. *Reinforcement Learning: An Introduction*, volume 1. MIT press Cambridge, 1998.

Van Dinh, Bui, Kim, Do Sang, and Jiao, Liguo. Convergence analysis of algorithms for DC programming. *arXiv preprint arXiv:1508.03899*, 2015.

Van Roy, Benjamin. Performance loss bounds for approximate value iteration with state aggregation. *Mathematics of Operations Research*, 31(2):234–244, 2006.

Wang, Boxiang and Zou, Hui. Another look at distance-weighted discrimination. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 80(1):177–198, 2018.

Watson, GA. On matrix approximation problems with Ky Fan*k* norms. *Numerical Algorithms*, 5(5):263–272, 1993.

Wen, Bo, Chen, Xiaojun, and Pong, Ting Kei. A proximal difference-of-convex algorithm with extrapolation. *Computational Optimization and Applications*, pp. 1–28, 2017.

Yang, Lin F, Braverman, Vladimir, Zhao, Tuo, and Wang, Mengdi. Dynamic partition of complex networks. *arXiv preprint arXiv:1705.07881*, 2017.

Zhang, Anru and Wang, Mengdi. Optimal state compression of Markov processes via empirical low-rank estimation. *arXiv preprint arXiv:1802.02920*, 2018.