

Precision Medicine: Lecture 14

Applications of Deep Learning to Precision Medicine

Michael R. Kosorok,
Nikki L. B. Freeman and Owen E. Leete

Department of Biostatistics
Gillings School of Global Public Health
University of North Carolina at Chapel Hill

Fall, 2021

Outline

Sufficient Markov Decision Processes with Alternating Deep Neural Networks

Estimating Individualized Optimal Combination Therapies Through Outcome Weighted Deep Learning Algorithms

Introduction

- ▶ Sequential decision problems arise in many application areas
 - ▶ Autonomous Vehicles
 - ▶ Finance
 - ▶ Logistics
 - ▶ Robotics
 - ▶ Healthcare
- ▶ Markov decision processes (MDPs) are the primary mathematical model for sequential decision problems
- ▶ Almost any decision process can be made into an MDP
- ▶ Coercing a decision process into the MDP framework can lead to high-dimensional system information that is difficult to model

Motivation

- ▶ High-dimensional, infinite horizon MDPs can often be represented by low dimensional approximations
- ▶ Option 1: Create a finite discretization of the space and treat the process as a finite MDP
 - ▶ Can result in a significant loss of information
 - ▶ Can be difficult to apply when the system state information is continuous and high-dimensional
- ▶ Option 2: Construct a low-dimensional summary of the underlying system states
 - ▶ No guarantee low-dimensional summary contains salient features needed for making good decisions
- ▶ Can we find a good low dimensional approximation?

Setup and Notation

- ▶ The observed data are

$$\{(\mathbf{S}_i^1, A_i^1, U_i^1, \mathbf{S}_i^2, \dots, A_i^T, U_i^T, \mathbf{S}_i^{T+1})\}$$

- ▶ $T \in \mathbb{N}$ - Observation time
 - ▶ $\mathbf{S}^t \in \mathbb{R}^{p_t}$ - Summary of information until time t
 - ▶ $A^t \in \mathcal{A} = \{1, \dots, K\}$ - Decision made at time t
 - ▶ $U^t = U^t(\mathbf{S}^t, A^t, \mathbf{S}^{t+1})$ - Quantifies momentary “goodness” of current action / state transition
-
- ▶ We assume that U is bounded such that $\sup_t |U^t| \leq M$
 - ▶ The observed data has a time horizon T , but the method should work for any time horizon

Markov and Homogeneous

- ▶ We assume that the process is Markov and homogeneous in that

$$P(\mathbf{S}^{t+1} \in \mathcal{G}^{t+1} | A^t, \mathbf{S}^t, \dots, A^1, \mathbf{S}^1) = P(\mathbf{S}^{t+1} \in \mathcal{G}^{t+1} | A^t, \mathbf{S}^t),$$

where $\mathcal{G}^{t+1} \subseteq \text{dom} \mathbf{S}^{t+1}$, and the probability measure does not depend on t

- ▶ These conditions may not be satisfied without some modification
 - ▶ For any process $(\mathbf{S}^1, A^1, \mathbf{S}^2 \dots)$ we can define $\tilde{\mathbf{S}} = (\mathbf{S}^t, A^{t-1}, \dots, \mathbf{S}^{t-m_t})$ where m_t is chosen such that the Markov property holds
 - ▶ Augmenting the state with a variable for time, i.e. defining the state at time t to be $(\tilde{\mathbf{S}}^t, t)$, can ensure homogeneity

Decision Strategy

- ▶ The decision strategy $\pi : \mathcal{S} \mapsto \mathcal{A}$, makes decision $\pi(\mathbf{s}^t)$ when presented with $\mathbf{S}^t = \mathbf{s}^t$ at time t
- ▶ Let $\bar{\mathbf{a}}^t = (a^1, a^2, \dots, a^t)$ and $\bar{\mathbf{s}}^t = (\mathbf{s}^1, \mathbf{s}^2, \dots, \mathbf{s}^t)$ be the action and state histories at time t
- ▶ $\mathbf{S}^{*t}(\bar{\mathbf{a}}^{t-1})$ is the potential state under trajectory $\bar{\mathbf{a}}^{t-1}$
- ▶ The potential state under decision strategy π is

$$\mathbf{S}^{*t}(\pi) = \sum_{\bar{\mathbf{a}}^{t-1}} \mathbf{S}^{*t}(\bar{\mathbf{a}}^{t-1}) \prod_{v=1}^{t-1} 1_{\pi\{\mathbf{S}^{*v}(\bar{\mathbf{a}}^{v-1})\}=a^v}$$

Value Function

- ▶ The potential utility under decision strategy π is

$$U^{*t}(\pi) = U[\mathbf{S}^{*t}(\pi), \pi(\mathbf{S}^{*t}(\pi)), \mathbf{S}^{*(t+1)}(\pi)]$$

- ▶ The discounted mean utility under π is

$$V(\pi) = \mathbb{E} \left\{ \sum_{t \geq 1} \gamma^{t-1} U^{*t}(\pi) \right\}$$

- ▶ For a class of decision strategies Π , the optimal decision strategy $\pi^{\text{opt}} \in \Pi$ satisfies $V(\pi^{\text{opt}}) \geq V(\pi) \forall \pi \in \Pi$

Sufficient Markov Decision Processes

- ▶ It can be difficult to construct a high-quality estimator of π^{opt} when the states \mathbf{S}^t are high-dimensional
- ▶ For any map $\phi : \mathcal{S} \mapsto \mathbb{R}^q$, define $\mathbf{S}_\phi^t = \phi(\mathbf{S}_t)$
- ▶ ϕ induces a sufficient MDP for π^{opt} if $(\bar{\mathbf{A}}^t, \bar{\mathbf{S}}_\phi^{t+1}, \bar{\mathbf{U}}^t)$ contains all relevant information about π^{opt}
- ▶ Given a policy $\pi_\phi : \text{dom} \mathbf{S}_\phi^t \mapsto \mathcal{A}$, the potential utility under π_ϕ is

$$U_\phi^{*t}(\pi_\phi) = \sum_{\bar{\mathbf{a}}^t} U \left[\mathbf{s}^{*t}(\bar{\mathbf{a}}^{t-1}), a^t, \mathbf{s}^{*(t+1)}(\bar{\mathbf{a}}^t) \right] \prod_{v=1}^t \mathbf{1}_{\pi_\phi \{ \mathbf{s}_\phi^{*v}(\bar{\mathbf{a}}^{v-1}) \} = a^v}$$

Sufficient Markov Decision Processes

- ▶ Let $\Pi \subseteq \mathcal{A}^{\mathcal{S}}$ denote a class of decision strategies defined on \mathcal{S} and $\Pi_{\phi} \subseteq \mathcal{A}_{\phi}^{\mathcal{S}}$ a class of decision strategies defined on $\mathcal{S}_{\phi} = \text{dom} \mathbf{S}_{\phi}^t \subseteq \mathbb{R}^q$
- ▶ The pair (ϕ, Π_{ϕ}) induces a sufficient MDP for π^{opt} within Π if the following conditions hold for all $t \in \mathbb{N}$:
 - ▶ The process $(\bar{\mathbf{A}}^t, \bar{\mathbf{S}}_{\phi}^{t+1}, \bar{\mathbf{U}}^t)$ is Markov and homogeneous
 - ▶ There exists $\pi_{\phi}^{\text{opt}} \in \text{argmax}_{\pi \in \Pi_{\phi}} V(\pi)$ which can be written as $\pi^{\text{opt}} = \pi_{\phi}^{\text{opt}} \circ \phi$, where

$$\pi_{\phi}^{\text{opt}} \in \text{argmax}_{\pi_{\phi} \in \Pi_{\phi}} \mathbb{E} \left\{ \sum_{t \geq 1} \gamma^{t-1} U_{\phi}^{*t}(\pi_{\phi}) \right\}$$

- ▶ It suffices to store only the process $\{(\bar{\mathbf{S}}_{\phi,i}^{t+1}, \bar{\mathbf{A}}_i^t, \bar{\mathbf{U}}_i^t)\}_{i=1}^n$

Conditional Independence

- ▶ Verifiable conditions for checking that (ϕ, Π_ϕ) induces a sufficient MDP require a few assumptions
 - ▶ Consistency: $\mathbf{S}^t = \mathbf{S}^{*t}(\bar{\mathbf{A}}^{t-1})$
 - ▶ Positivity: $P(A^t = a^t | \bar{\mathbf{S}}^t = \bar{\mathbf{s}}^t, \bar{\mathbf{A}}^{t-1} = \bar{\mathbf{a}}^{t-1}) > 0$
 - ▶ Sequential ignorability: $\{\mathbf{S}^{*t}(\bar{\mathbf{a}}^{t-1})\}_{t \geq 1} \perp A^t | \bar{\mathbf{S}}^t, \bar{\mathbf{A}}^{t-1}$
- ▶ Define $\mathbf{Y}^{t+1} = \{U^t, (\mathbf{S}^{t+1})^T\}^T$ for all $t \in \mathbb{N}$
- ▶ Let $(\mathbf{S}^1, A^1, U^1, \mathbf{S}^2, \dots)$ be an MDP that satisfies the above assumptions. Suppose that there exists $\phi : \mathcal{S} \mapsto \mathbb{R}^q$ such that

$$\mathbf{Y}^{t+1} \perp \mathbf{S}^t | \mathbf{S}_\phi^t, A^t$$

then, (ϕ, Π_ϕ) induces a sufficient MDP for π^{opt} within Π

Conditional Independence

- ▶ Let $(\mathbf{S}^1, A^1, U^1, \mathbf{S}^2, \dots)$ be an MDP that satisfies the previous assumptions. Suppose that there exists $\phi : \mathcal{S} \mapsto \mathbb{R}^q$ such that at least one of the following conditions hold:

(i) $\{\mathbf{Y}^{t+1} - \mathbb{E}(\mathbf{Y}^{t+1} | \mathbf{S}_\phi^t, A^t)\} \perp \mathbf{S}^t | A^t$

(ii) $\{\mathbf{S}^t - \mathbb{E}(\mathbf{S}^t | \mathbf{S}_\phi^t)\} \perp (\mathbf{Y}^{t+1}, \mathbf{S}_\phi^t) | A^t$

then $\mathbf{Y}^{t+1} \perp \mathbf{S}^t | \mathbf{S}_\phi^t, A^t$

- ▶ This result can be used to verify the conditional independence condition using Brownian distance covariance

Variable Screening

- ▶ Conditional independence may be too strong of an assumption in the presence of certain noise variables
- ▶ For example, let $\{\mathbf{B}^t\}_{t \geq 1}$ denote a homogeneous Markov process independent of $(\mathbf{S}^1, A^1, U^1, \mathbf{S}^2, \dots)$
- ▶ \mathbf{Y}^{t+1} need not be conditionally independent of $\{(\mathbf{S}^t)^T, (\mathbf{B}^t)^T\}^T$ given \mathbf{S}^t , but π^{opt} does not depend on $\{\mathbf{B}^t\}_{t \geq 1}$
- ▶ For any map $\phi : \mathcal{S} \mapsto \mathbb{R}^q$, let $\mathbf{Y}_\phi^{t+1} = \{U^t, (\mathbf{S}_\phi^{t+1})^T\}^T$
- ▶ If $\mathbf{Y}_\phi^{t+1} \perp \mathbf{S}^t | \mathbf{S}_\phi^t, A^t$ then, (ϕ, Π_ϕ) induces a sufficient MDP for π^{opt} within Π

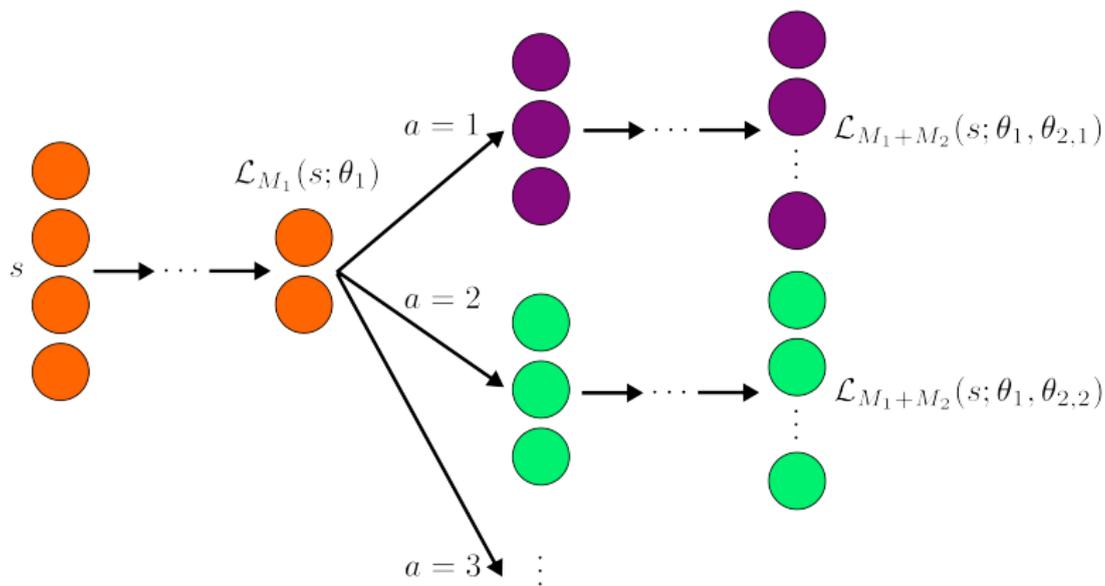
Alternating Deep Neural Networks

- ▶ Let $\mathcal{S} = \mathbb{R}^P$ and consider summary functions $\phi : \mathbb{R}^P \mapsto \mathbb{R}^q$ that are representable as multi-layer neural networks
- ▶ To construct a data-driven summary function ϕ , we require a model for the regression of \mathbf{Y}^{t+1} on \mathbf{S}_ϕ^t and \mathbf{A}^t
 - ▶ This predictive model can also be representable as a multi-layer neural network
- ▶ The model can be visualized as two connected multi-layer neural networks
 - ▶ One that composes the feature map ϕ
 - ▶ One that models the regression of \mathbf{Y}^{t+1} on \mathbf{S}_ϕ^t and \mathbf{A}^t

Alternating Deep Neural Networks

- ▶ The First M_1 layers of of the NN determine the feature map $\phi \equiv \mathcal{L}_{M_1}(\mathbf{s}, \theta_1)$
- ▶ The following M_2 layers fit the regression model $\mathcal{L}_{M_2}(\mathbf{s}, \theta_1, \theta_2)$
 - ▶ Separate regression models are fit for each action a
- ▶ There are several tuning parameters including the number of layers, width of the layers, and the dimension of the feature map $\phi : \mathcal{S} \mapsto \mathbb{R}^q$
- ▶ The dimension q is chosen to be the lowest dimension for which the Brownian distance covariance test of independence fails to reject at a pre-specified error level

Alternating Deep Neural Networks



Conclusions

- ▶ Choosing a parsimonious representation of a decision process that fits the MDP model is non-trivial
- ▶ The deep neural networks can model complex and nonlinear structure in the high-dimensional state space
- ▶ The method proposed above can effectively reduce the dimension of the state space without adversely impacting the performance
- ▶ Possible future work includes
 - ▶ Online estimation of the feature map
 - ▶ States with complex data structures (e.g. images and text)

Outline

Sufficient Markov Decision Processes with Alternating Deep Neural Networks

Estimating Individualized Optimal Combination Therapies Through Outcome Weighted Deep Learning Algorithms

Introduction

- ▶ Many conditions have multiple treatments available
- ▶ Combination therapies are becoming increasingly common
- ▶ Multiclass individualized treatment rule (ITR) estimation methods are not scalable with large numbers of treatment options
- ▶ Existing algorithms may not be adequate for finding ITRs with combination therapies

Motivation

- ▶ Patients with type 2 diabetes often receive multiple medications because a single treatment may be insufficient to effectively control the blood glucose level
- ▶ Medications exerting their effects through different mechanisms can increase effectiveness of the therapy
 - ▶ Sulfonylurea (SU) increases insulin release from β -cells in the pancreas
 - ▶ DDP4 increases incretin level, which inhibits glucagon release
- ▶ Effects may be additive or weakly interacting because DDP4 and SU function through different biological pathways

Multiclass vs Multilabel

- ▶ Combination therapies often allow patients to be assigned to any combination of K possible treatments
- ▶ Multiclass classification problems would have 2^K classes
 - ▶ Dimension increases exponentially with the number of treatments
 - ▶ Proposed strategies to simplify computation have many drawbacks
- ▶ Multilabel classification considers K binary treatment choices (yes/no for each drug)
 - ▶ Reduces computational cost of estimation
 - ▶ Requires treatment effects to be additive or small interactions

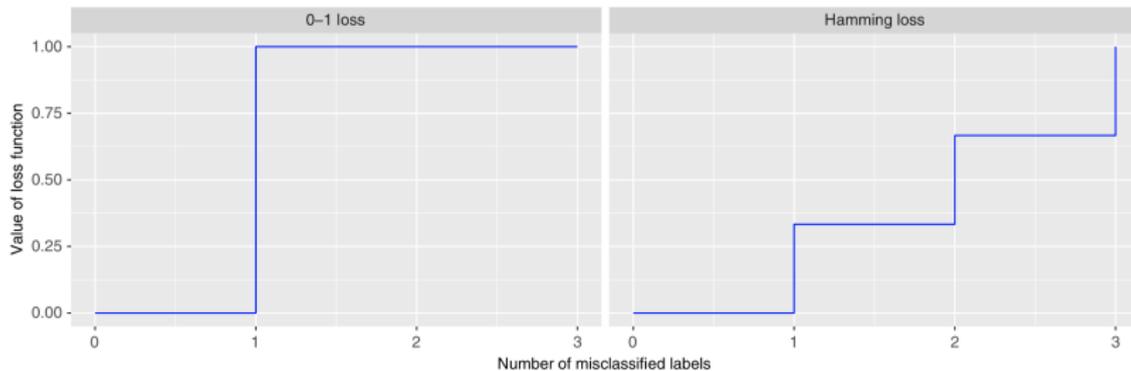
Hamming Loss

- ▶ Let $A = (A_1, \dots, A_K) \in \mathcal{A}$ represent a vector of length K , where $\mathcal{A} = \{-1, 1\}^K$
- ▶ Accordingly, the decision rule is $D(X) = (D_1(X), \dots, D_K(X)) \in \{-1, 1\}^K$
- ▶ A commonly used loss function for multilabel classification is the Hamming loss

$$\frac{1}{K} \sum_{k=1}^K I\{A_k \neq D_k(X)\}$$

- ▶ The Hamming loss quantifies the proportion of mislabeled treatments

Hamming Loss



- ▶ Like the 0-1 loss, the Hamming loss is discontinuous and difficult to optimize
- ▶ To aid computation we will use a convex surrogate loss

OWL with Multilabel Classification

- ▶ Consider the following loss function for a given decision rule $D(X)$:

$$L(D) = \frac{1}{n} \sum_{i=1}^n \frac{R_i}{\pi_{A_i}} \frac{1}{K} \sum_{k=1}^K I\{A_k \neq D_k(X)\}$$

where $\pi_{A_i} = Pr(A = a|X)$

- ▶ For multiclass classification, this loss reduces to the outcome weighted 0-1 loss
- ▶ In observational studies π_{A_i} can be difficult to estimate when K is large. Possible approaches:
 - ▶ $P(A = a|X) = P(A = a)$
 - ▶ $P(A = a|X) = \prod_{k=1}^K P(A_k = a_k|X)$

Decision Rule with Deep Learning

- ▶ Any suitable classifier can fit into the proposed framework
- ▶ We consider neural networks because of their advantages in shared subspace and scalable computation
- ▶ Let $\widehat{D}(X) = (\widehat{D}_1(X), \dots, \widehat{D}_K(X))$ be the estimated decision rule
- ▶ Using the hinge loss, we can formulate the optimization problem as

$$\min_{\theta} \frac{1}{n} \sum_{i=1}^n \frac{R_i}{\pi_{A_i}} \frac{1}{K} \sum_{k=1}^K [1 - A_{ik} \widehat{D}_k(X_i)]_+$$

where θ represents all parameters in the NN

Overfitting

- ▶ Due to strong representative power of DNN, overfitting can be an issue
- ▶ The authors recommend a regularization strategy with penalization by solving the following problem

$$\min_{\theta} \frac{1}{n} \sum_{i=1}^n \frac{R_i}{\pi_{A_i}} \frac{1}{K} \sum_{k=1}^K [1 - A_{ik} \hat{D}_k(X_i)]_+ + \lambda f(\theta)$$

- ▶ Let L be the number of layers and W_{ℓ} be the weight matrix from the ℓ^{th} layer
 - ▶ Ridge penalty: $f(\theta) = \sum_{\ell=1}^L \|W_{\ell}\|_F^2$ (Frobenius norm)
 - ▶ Lasso penalty: $f(\theta) = \sum_{\ell=1}^L \|W_{\ell}\|_1$
- ▶ Other methods to prevent overfitting can also be used

Fisher Consistency of Hamming loss

- ▶ Define the risk of the outcome weighted 0-1 loss as

$$\mathcal{R}(D) = \mathbb{E} \left[\frac{R}{\pi_A} I\{A \neq D(X)\} \right]$$

- ▶ And the risk of the outcome weighted Hamming loss as

$$\mathcal{R}_H(D) = \mathbb{E} \left[\frac{R}{\pi_A} \frac{1}{K} \sum_{k=1}^K I\{A_k \neq D_k(X)\} \right]$$

- ▶ Any decision rule \tilde{D} such that $\mathcal{R}_H(\tilde{D}) = \inf_D \{\mathcal{R}_H(D)\}$ also satisfies $\mathcal{R}(\tilde{D}) = \inf_D \{\mathcal{R}(D)\}$
- ▶ Since the outcome weighted 0-1 loss is fisher consistent, so is the outcome weighted Hamming loss

Multilabel Consistency of the Surrogate Loss

- ▶ Define the risk of the surrogate outcome weighted Hamming loss as

$$\Phi_H(D) = \mathbb{E} \left[\frac{R}{\pi_A} \frac{1}{K} \sum_{k=1}^K \phi(A_k D_k(X)) \right]$$

where ϕ is a predefined convex function

- ▶ The surrogate loss is consistent with outcome weighted Hamming loss if ϕ is one of the following
 1. Exponential: $\phi(x) = e^{-x}$
 2. Hinge: $\phi(x) = (1 - x)_+$
 3. Least squares: $\phi(x) = (1 - x)^2$
 4. Logistic regression: $\phi(x) = \ln(1 + e^{-x})$

Conclusions

- ▶ Using multilabel classification can make finding ITRs for combination therapies tractable
- ▶ The proposed Hamming loss does not account for strong interactions among treatments which may limit its applicability
- ▶ A generalization of the Hamming loss can handle interactions, but the theoretical properties are unknown
- ▶ Possible extensions include multiple decision points and estimating contrasts of treatment effects in combination therapies
- ▶ Multimodal propensity scores (that address interactions between treatment modes)