Operations Research Letters 44 (2016) 394–399

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Robust response-guided dosing

Saumya Sinha [a], Jakob Kotas [a], Archis Ghate [b,*]

[a] Department of Applied Mathematics, University of Washington, Seattle, WA, USA

[b] Industrial and Systems Engineering, Box 352650, University of Washington, Seattle, WA 98195, USA

ARTICLE INFO

Article history:
Received 4 October 2015
Received in revised form
23 March 2016
Accepted 23 March 2016
Available online 1 April 2016

Keywords:
Stochastic dose–response
Robust Markov decision processes
Medical treatment planning

ABSTRACT

In response-guided dosing (RGD), doses are adapted to the uncertain progression of each patient's disease condition. A stochastic dynamic program was recently developed for RGD. We study its robust counterpart, where the dose–response distribution belongs to an uncertainty set. For interval uncertainty sets, we prove that it is optimal to administer higher doses in worsening disease. When a certain scaling of a nominal distribution describes the interval, optimal doses also increase in the scaling parameter. Theory is illustrated via numerical results.

1. Background and motivation

Many diseases are treated with therapeutic drugs over multiple sessions where dosing guidelines follow a one-size-fits-all philosophy. For instance, a daily radiation dose of 2 Gy for 35 days is recommended for head-and-neck cancer [7], and a weekly dose of 180 μg interferon for 48 weeks is prescribed for hepatitis C [9]. While such guidelines are easy to implement, their potential drawbacks include inadequate disease control, over- and under-dosing, and unsatisfactory cost-effectiveness [4]. Response-guided dosing (RGD) is an alternative paradigm, where doses are adapted to the observed uncertain evolution of a patient's disease condition. RGD attempts to balance a fundamental trade-off in medicine — higher doses achieve better disease-control at the cost of undesirable side-effects, whereas lower doses sacrifice disease-control in favor of lesser adverse effects.

Most clinical implementations of RGD are, however, somewhat *ad hoc*. Indeed, Murphy et al. [8] have commented: "*despite the activity in evaluating adaptive treatment strategies, the development of data collection and analytic methods that directly inform the construction of adaptive treatment strategies lags behind.*" To address this concern, a stochastic dynamic programming (DP) framework for choosing doses in RGD was recently proposed by our research group in Kotas and Ghate [6]. The idea there was to balance the total disutility of doses delivered over the treatment course against the disease condition reached by the end of the treatment course. The evolution of this disease condition was characterized using a stochastic dose–response model — a formula that yields the next disease condition as a function of the current disease condition and of the dose used in the current treatment session, subject to uncertainty in a dose–response parameter. Kotas and Ghate proved, under natural convexity and monotonicity assumptions on this formula and on the disutility functions, that optimal doses are increasing in worsening disease conditions.

One limitation of this stochastic DP is that the decision-maker is assumed to know at the outset the probability mass function (pmf) of the dose–response parameter. Any *a priori* estimate of this pmf, however, is subject to estimation errors. To tackle the resulting ambiguity, we present here a robust counterpart of the Kotas and Ghate model (henceforth called the "nominal" model).

In our robust formulation, the pmf of the dose–response parameter will be assumed to belong to an uncertainty set. Uncertainty sets are often composed of pmfs that are in some sense "close to" a nominal pmf, which may have been estimated *a priori* from a clinical trial [1,5,10]. Roughly speaking, the decision-maker then follows a conservative approach whereby he/she attempts to find a dosing policy that minimizes the worst-case expected disutility over all pmfs from this uncertainty set. Examples of uncertainty sets include the interval set, the maximum likelihood set, the relative entropy set, and the ellipsoidal set. Although our general robust RGD model accommodates any of these sets, we illustrate our results in detail using the interval model. We show that the so-called inner maximization problem in the Bellman's equations for robust RGD with the interval uncertainty set is a linear program (LP) that can be solved analytically. Moreover, an optimal solution to this inner problem, that is, the worst-case pmf, does not depend on the observed disease condition

* Corresponding author.
E-mail address: archis@uw.edu (A. Ghate).

http://dx.doi.org/10.1016/j.orl.2016.03.012
0167-6377/

and the dose chosen. This in turn implies that there exists a monotone dosing policy that is optimal to the robust stochastic DP, thus extending the main theoretical result from the nominal model of Kotas and Ghate (see Section 3). This extension is not only of theoretical interest but also significantly simplifies the computation of our robust optimal dosing policy. In particular, we show that the state–action invariant structure of the worst-case pmf makes the robust problem only as hard to solve as the nominal problem. We further analyze in Section 4 a specific and common single-parameter formulation of the interval uncertainty set and provide a simple condition on the dose–response formula under which optimal doses vary monotonically with this parameter. We conclude by presenting numerical results on a hypothetical disease with an inverse-power dose–response function.

2. Review of the nominal model

Let T denote the number of sessions, indexed by $t = 1, 2, \ldots, T$, in a treatment course. At the beginning of each session, the decision-maker observes a numerical score of the patient's disease condition, and chooses a dose for that session. These numerical scores belong to a compact interval $X \subseteq \mathbb{R}$. Smaller numbers in this set represent less severe disease. The disease condition at the beginning of session t is denoted by $x_t \in X$. The dose level chosen by the decision-maker for this session after observing x_t is denoted by d_t. Dose levels d_t belong to the interval $D \triangleq [0, \bar{d}] \subset \mathbb{R}$, where $\bar{d}$ is a finite upper bound on permissible dose levels.

For $t = 1, 2, \ldots, T$, disease conditions evolve according to dynamics $x_{t+1} = x_t + f(d_t; \Theta_t)$, for $x_t, x_{t+1} \in X$ and $d_t \in D$. All standard dose–response functions such as linear, Michaelis–Menten, inverse-power, Emax, Hill's, exponential, exponential linear–quadratic, power law, Gompertz, and Beta-Poisson can be expressed in this form (after taking logarithms in some cases). Please see Kotas and Ghate and references therein for detailed descriptions of these functions. Here, Θ_t are independent and identically distributed dose–response parameters in sessions t. Independence across sessions is somewhat restrictive although common in the literature (see, for example, Chapter 4 of [3], and also [11]). Kotas and Ghate employed this assumption in their nominal model as well, and it holds when consecutive sessions are "sufficiently separated" from a biochemical viewpoint. Random variables Θ_t take values from a finite set $\Omega \triangleq \{\theta_1, \theta_2, \ldots, \theta_n\}$. Their pmf, denoted by $p(\cdot)$, is known to the decision-maker. We assume that the function $f(\cdot; \theta)$ is continuous over D for each $\theta \in \Omega$.

Aversion to dose is modeled using a continuous disutility function $c : D \to \mathbb{R}_+$. Since D is compact, continuity of $c(\cdot)$ implies that it is bounded. Examples include linear, quadratic, and exponential functions. Aversion to disease conditions x_{T+1} at the end of the treatment course is modeled using a continuous and bounded disutility function $h : X \to \mathbb{R}_+$. Examples include linear, quadratic, exponential, and ramp (where the disutility is zero up to a disease-condition threshold and grows linearly thereafter).

Let $J_t(x_t)$ denote the minimum total expected disutility accumulated by the end of the treatment course, given that the disease condition at the beginning of the tth session is x_t. These optimal cost-to-go functions $J_t(\cdot)$ are unique solutions of Bellman's equations

$$J_t(x_t) = \min_{d_t \in D} \left\{ c(d_t) + \sum_{\theta \in \Omega} J_{t+1}(x_t + f(d_t; \theta)) p(\theta) \right\},$$
$$\forall x_t \in X, \text{ and } t = 1, 2, \ldots, T, \tag{1}$$

with the boundary condition $J_{T+1}(x) = h(x)$ for all $x \in X$. Problem (1) involves optimizing a continuous function over the nonempty compact set D and hence it has an optimal solution. Doses that attain the above minima define an optimal RGD policy. Bellman's equations (1) can easily be solved approximately using discretization of X and D. As such, the nominal problem is computationally tractable. Moreover, it is shown in Kotas and Ghate, under two assumptions (stated below) on the dose–response function and on the disutility functions, that in each treatment session there exist optimal doses that increase as the disease condition worsens.

Assumption 2.1 (*Monotone and Convex Dose–Response*)**.** The function $f(\cdot; \theta)$ is decreasing and convex in dose over D for every $\theta \in \Omega$.

Assumption 2.2 (*Increasing and Convex Disutilities*)**.** The disutility function $c(\cdot)$ is increasing and convex over D; the disutility function $h(\cdot)$ is increasing and convex over X.

A detailed justification for these assumptions was provided by Kotas and Ghate. In particular, Assumption 2.2 encodes a risk-averse decision-maker. Several examples of clinically relevant functions that satisfy these assumptions were also listed by Kotas and Ghate; as such, we do not believe these assumptions to be particularly restrictive.

Theorem 2.3 (*Theorem 5.3 in Kotas and Ghate*)**.** *Under the above assumptions, optimal dose levels increase with worsening disease conditions in each treatment session.*

3. The robust stochastic DP

As opposed to the nominal problem, we now consider the case where the pmf of the dose–response parameter is not known to the decision-maker. Pursuing standard practice in robust stochastic DP [1,5,10], we assume that the pmf of Θ_t is only known to lie in some set $\mathcal{P}$. In the robust optimization parlance, set $\mathcal{P}$ is called the uncertainty set and is composed of all "plausible" pmfs. This set is often chosen so that it includes all pmfs that are "close to" some nominal pmf. More precisely, let $\Delta \triangleq \{p(\cdot) : p(\theta) \geq 0, \ \sum_{\theta \in \Omega} p(\theta) = 1\}$ be the probability simplex in $\mathbb{R}^n$, and let $\mathcal{P} \subseteq \Delta$. Then, the worst-case total expected disutility is minimized by solving the robust Bellman's equations

$$\tilde{J}_t(x_t) = \min_{d_t \in D} \left\{ \overbrace{\max_{p_t(\cdot) \in \mathcal{P}} \left(c(d_t) + \sum_{\theta \in \Omega} \tilde{J}_{t+1}(x_t + f(d_t; \theta)) p_t(\theta) \right)}^{\text{"inner problem"}} \right\},$$
$$\forall x_t \in X, \text{ and } t = 1, 2, \ldots, T, \tag{2}$$

with the boundary condition $\tilde{J}_{T+1}(x) = h(x)$ for all $x \in X$. Here, $\tilde{J}_t(\cdot)$, for $t = 1, 2, \ldots, T+1$, are called the robust optimal cost-to-go functions, and an optimal solution to the inner problem is called a worst-case distribution.

The robust stochastic DP is computationally tractable if the inner maximization problem is easy to solve. This occurs, for example, when $\mathcal{P}$ is chosen to be a convex set. This, combined with the linearity (in $p_t(\cdot)$) of the objective function, implies that the inner problem is convex. Some examples of convex uncertainty sets are interval, maximum likelihood, ellipsoidal and entropy [1,5,10]. We focus on the interval uncertainty model in the subsequent discussion.

The interval model is motivated by statistical estimates of confidence intervals on the pmf components. It can also be obtained by projecting the ellipsoidal or maximum likelihood uncertainty sets onto the coordinate axes [1,10]. In this model [1], the uncertainty set $\mathcal{P}$ is defined such that the probabilities $p(\theta)$, for $\theta \in \Omega$, belong to an interval. More precisely, $\mathcal{P} \triangleq \{p(\theta) \in \Delta : p_L(\theta) \leq p(\theta) \leq p_H(\theta)\}$, for some constants $p_L(\theta)$ and