Stochastic Analysis of Therapeutic Modalities Using a Database of Patient Responses

D.S. Bayard [1] A. Botnan [2] W.C. Shoemaker[3] and R. Jelliffe[4]
Laboratory of Applied Pharmacokinetics
University of Southern California
School of Medicine
Los Angeles, CA 90033

Topical Areas:

- Patient Informatics
- Decision Support Systems
- Software Systems in Medicine

[1]Senior Research Scientist, Jet Propulsion Laboratory, Pasadena CA, 91109
[2]Research Associate, Laboratory of Applied Pharmacokinetics, USC School of Medicine
[3]Professor of Surgery and Anesthesia, Departments of Surgery and Anesthesia, USC School of Medicine
[4]Corresponding author, Professor of Medicine, Director Laboratory of Applied Pharmacokinetics, USC School of Medicine, 2250 Alcazar St. Los Angeles CA 90033, USA, Tel (323) 442-1300, Fax (323) 442-1302, *jelliffe@hsc.usc.edu*

Stochastic Analysis of Therapeutic Modalities Using a Database of Patient Responses

This paper proposes a new method for stochastic analysis and control which does not require a model, but rather is constructed directly from a raw database of patient responses to therapy. Roughly speaking, the basic idea is to evaluate a control (a therapeutic policy or modality) which has, on the average, proved to work well for similar patients in the database. By "similar" is meant patients who have the same covariates and who are in similar dynamical states. These concepts will be made more precise in the paper. The proposed stochastic analysis and control approach for databases is new, although it is motivated by methods of machine learning put forth in [1][2] and methods of dynamic programming for stochastic control given in [3][4].

Concept of State

The first key step is to develop the definition of a "state" for a given patient in the database at a given stage of treatment. Many definitions of state are possible. Since we are working with a database of measurements, it is useful to define the state vector directly in terms of the available measurement information. Assume there are L different types of measurements obtained on a given patient (e.g., cardiac index, blood pressure, pulse oximetry, transcutaneous O_2 and CO_2 tensions, etc.). Then for each measurement type, denoted as y_ℓ, define the state vector as a concatenation of the value y_ℓ itself, with its first and second time derivatives y'_ℓ, y''_ℓ, and with its first integral $\int y_\ell \; dt$, as follows,

$$x(t_k) = \left[y_1(t_k), y'_1(t_k), y''_1(t_k), \int_0^{t_k} y_1 \; dt, ..., y_L(t_k), y'_L(t_k), y''_L(t_k), \int_0^{t_k} y_L \; dt \right] \tag{1}$$

Implicit Dynamic Model

It is convenient to think of the propagation of the patient's state x_k at time t_k, to his state x_{k+1} at time t_{k+1} as obeying the following nonlinear dynamical system with process noise w_k, parameter vector p, and control u_i i.e.,

$$x_{k+1} = f(x_k, u_i, p, w_k) \tag{2}$$

For simplicity, p is assumed to be drawn from a finite set formed by enumerating all useful combinations of covariates,

$$p \in \{p_1, ..., p_M\} \tag{3}$$

Both the parameter vector p of covariates and process noise w help to explain the variability of responses seen in the database. Specifically, the covariates help to distinguish gross differences in responses due to major categories of complications (e.g., with vs. without head injuries, blunt vs. penetrating trauma, truncal vs. nontruncal trauma, age stratifications, etc.) On the other hand, process noise helps to explain smaller differences between patients with identical covariates but with different responses to the same therapy.

Control Inputs

It will be assumed that there is only a finite set of M control inputs in (2) that can be applied to the system. Specifically, the control $u(t_k)$ at time t_k is assumed to be drawn from the finite set,

$$u(t_k) \in \{u_1, ..., u_M\} \tag{4}$$

where u_i are controls (therapeutic modalities) such as fluid therapy with crystalloids, with colloids, with whole blood or packed RBC's, or various vasoactive drugs such as dobutamine, dopamine, and other possible modalities.

Nearest Neighbors

Once the state vector and covariates are defined, the key concept of a "nearest neighbor" can be put forth. Given the state x_k and a certain covariate p, the N nearest-neighbor states (denoted as $\{x_k^j\}_{j=1}^N = \mathcal{N}(x_k, p)$) are defined as the states x that are closest to x_k in the database and which share the same covariate vector p. Here, a measure of "closeness" is conveniently defined in terms of the quadratic distance between the state and its neighbors,

$$d(x, x_k) \triangleq (x_k - x)^T W (x_k - x) \tag{5}$$

where W is an appropriately chose weighting matrix.

Performance Measure

The performance measure is the probability of survival. For a patient in a given state x with covariate vector p, the survival probability is denoted by $S(x, p)$. It is evaluated simply by extracting the N nearest neighbor states to x of patients with the same covariate vector p and by noting the fraction of them that survived. For example, if N_s of the N nearest-neighbor states to x survived, then the survival probability is given as,

$$S(x, p) = \frac{N_s}{N} \tag{6}$$

Applications

Data Mining for Significant Events

One can now trace through individual patient histories in the database and compute the associated probability of survival $S(x, p)$ as a function of time. A large change, either up or down, indicates a potentially significant role of the associated therapeutic modality, and its underlying physiologic mechanisms. The entire database can be mined for therapeutic modalities having the most efficacy or harm under specific conditions using this technique.

Real-Time Diagnostic Tool

The stochastic analysis is not restricted to subjects contained within the database. In addition, it can be used on a real patient during his/her clinical care. In this case, the survival probability $S(x, u)$ is computed (by tracking the subject's nearest neighbors in the database) and provided to the clinician as a real-time feedback signal which indicates the condition of the patient as a function of time, and the efficacy of any therapeutic modalities administered.

Stochastic Control for Choosing the Optimal Therapy

By classifying a subject's nearest neighbors by the therapeutic modalities employed, the survival probabilities $P(u_i, x_k, p)$ can be computed separately for each modality u_i. A useful method for stochastic control then follows by administering the therapy associated with the highest value of $P(u_i, x_k, p)$. Intuitively, this therapeutic modality is optimal (and a worthwhile choice in practice) in the sense that it has the highest survival probability among nearest neighbors in the database.

The stochastic analysis and control approach is summarized in Figure 1. Examples of the applications discussed above will be included in the final paper, as applied to a large hemodynamic database of acutely ill patients [5].

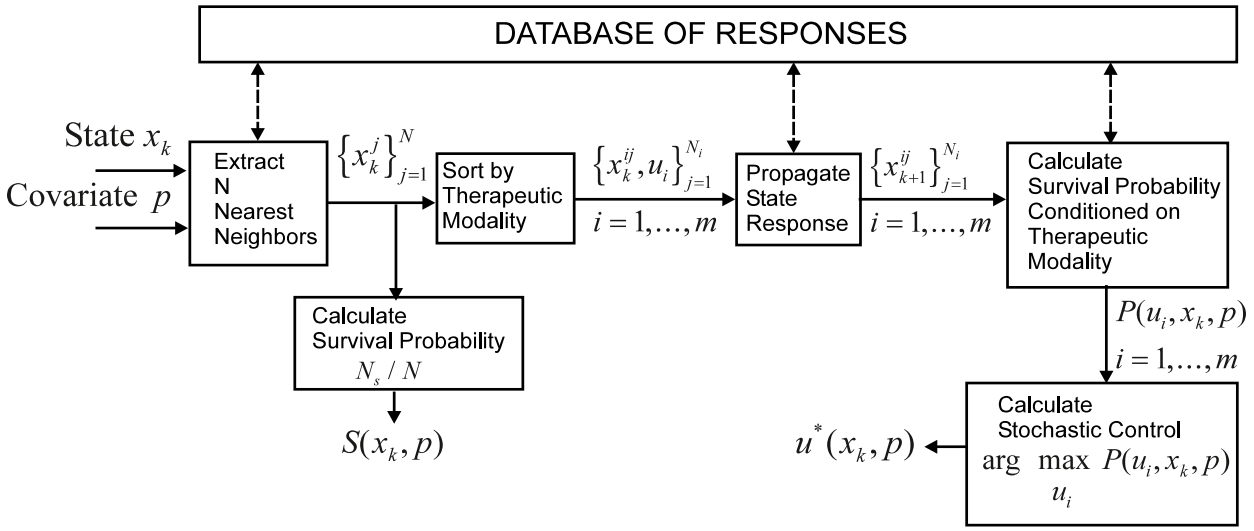

Figure 1: Stochastic analysis and control synthesized using a database of responses

Supported by NIH grant LM05401.

References

[1] R.S. Sutton and A.G. Barto, *Reinforcement Learning.* The MIT Press, Cambridge, Mass., 1998.

[2] D.P. Bertsekas and J.N. Tsitsiklis, *Neuro-Dynamic Programming.* Athena Scientific, Belmont, Mass. 1996.

[3] D.S. Bayard, "A Forward Method for Optimal Stochastic Nonlinear and Adaptive Control," IEEE Transactions on Automatic Control, Vol. 36, No. 9, pp. 1046-1053, Sept. 1991.

[4] D.S. Bayard, "Reduced Complexity Dynamic Programming Based On Policy Iteration," J. Mathematical Analysis and Applications, Vol. 170, No. 1, pp. 75-103, October 1992.

[5] W.C. Shoemaker, P.L. Appel and H.B. Kram, "Hemodynamic and oxygen transport responses in survivors and nonsurvivors of high risk surgery," Critical Care Medicine, Vol. 21, pp. 977-990, 1993.