Massachusetts Institute of Technology

Department of Electrical Engineering and Computer Science

6.242, Fall 2004: MODEL REDUCTION [*]

Proper Orthogonal Decomposition[1]

Here we discuss the method of model reduction via *proper orthogonal decomposition* (POD).

Motivation from statistics

POD is a method originating in statistical analysis of vector data. Consider the case when the same phenomena is measured m times, each measurement x_k, $k = 1, \ldots, m$ being a vector containing a large number n of real entries: $x_k \in \mathbf{R}^n$. (For example x_k could be a digital image obtained at the k-th experiment.) An important objective of statistical analysis of data is to discover interdependencies within the data, and to reduce the data set to a much smaller number $r \ll n$ of parameters.

Mathematically, the situation can be described by the optimization problem

$$\mathbf{E}|x - Px|^2 \mapsto \min,$$

where x is a vector random variable taking values in $\mathbf{R}^n$, $\mathbf{E}$ denotes the expected value, and P is a projection operator of rank r, i.e.

$$P = VU, \;\; UV = I_r.$$

The problem has an unsurprising solution in terms of r dominant eigenvectors of the covariance matrix

$$W = W_x = \mathbf{E}xx'.$$

---

[1] Version of October 6, 2004

If $\sigma_1^2 \geq \sigma_2^2 \geq \ldots$ are the ordered eigenvalues of W, $(\sigma_k \geq 0)$, and v_k are the corresponding orthonormalized eigenvectors, i.e.

$$W v_k = \sigma_k^2 v_k, \quad v_i' v_k = \delta_{ik},$$

then an optimal P is the orthonormal projection

$$P = VV', \quad V = [v_1 \quad v_2 \quad \ldots \quad v_r].$$

Moreover, the optimal P is unique if and only if $\sigma_r > \sigma_{r+1}$.

In statistical applications, the covariance matrix W_x is replaced by its estimate

$$\tilde{W}_x = \frac{1}{m} \sum_{k=1}^{m} x_k x_k' = \frac{1}{m} XX',$$

where x_k are the vector measurements, assumed to be *independent*, and X is the n-by-m matrix with columns x_k. Then σ_i are the singular numbers of $X/sqrt(m)$, and v_i are the corresponding singular vectors. Thus, there is no need to solve an eigenvalue problem for an n-by-n matrix $\tilde{W}_x$: instead, eigenmodes of the m-by-m matrix $X'X$ are to be found, where $m \ll n$.

POD for dynamical systems

Application of POD to dynamical system model reduction calls for using system state response in place of data vectors x_k. In the standard version, $x_k = x(t_k)$ are time samples of the full state response, obtained, usually, via simulation. Alternatively, for LTI state space models, the use of $x_k = (s_k I_n - A)^{-1} B$ is suggested. The resulting projection P is used to generate a reduced model via a standard projection scheme.

There are several serious objections to the use of POD for model reduction.

(a) When generated by dynamical system equations, the samples $x_k = x(t_k)$ or $x_k = (s_k I_n - A)^{-1} B$ are unlikely to be statistically independent. Hence, the statistical motivation does not really apply.

(b) Good fitting of the set $\{x_k\}$ by a low dimensional subspace does not mean good dynamical system approximation.

(c) In general, good projections for model reduction are not necessarily orthogonal.

Typically, an attempt to fix problems (a) and (b) leads to making $\tilde{W}_x$ an approximation of the controllability Gramian of the system. For example, a numerical integration quadrature

$$\int_0^\infty h(\theta)d\theta \approx \sum_{k=1}^m \rho_k^2 h(\theta_k),$$

where $\rho_k \geq 0$ are some coefficients, would lead to

$$x_k = x(t_k)\rho_k, \quad t_k = \theta_k,$$

where $x = x(t)$ is an approximation of zero input system response to initial condition $x(0) = B$, or, alternatively, to

$$x_k = (j\omega_k I_n - A)^{-1}B\rho_k, \quad \omega_k = \theta_k.$$

An attempt to fix problem (c) leads to working with two matrices: $\tilde{W}_x$ defined as before, and the dual $\tilde{W}_p$ defined by A', C' replacing A, B. In any case, the whole approach reduces to attempts to approximate the controllability and observability Gramians, and encounters serious difficulty in establishing quality guarantees.

Despite the objections listed above, model reduction via POD is quite popular, since the time domain samples $x_k = x(t_k)$ are easy to obtain whenever a numerical simulator for the system (linear or nonlinear) is available.

An example

Consider system

$$\begin{aligned} \dot{x}_1 &= -2x_1 + f, \\ \dot{x}_2 &= -x_2 + ax_1, \end{aligned}$$

where $a > 3$ is a parameter. Here

$$A = \begin{bmatrix} -2 & 0 \\ a & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

For the solution $x = x(t)$ of system equations with $x(0) = B$, we have

$$\bar{W}(t) = \int_0^T x(t)x(t)'dt = \begin{bmatrix} \alpha(t) & \beta(t) \\ \beta(t) & \gamma(t) \end{bmatrix},$$

where α, β, γ are non-negative numbers such that $\alpha(t) > \gamma(t)$ for small $t > 0$, and $\alpha(t) < \gamma(t)$ for large $t > 0$. Hence $\alpha(t_0) = \gamma(t_0)$ for some $t_0 > 0$, which means that the dominant normalized eigenvector V of $\bar{W}(t_0)$ is

$$V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Hence the projected system has an *unstable* pole

$$s = V'AV = \frac{a - 3}{2} > 0.$$