Singular Value Decomposition

Low-dimensional group

15 February 2005

Abstract

Introduction, by Juan Restrepo, to singular value decomposition in the context of data assimilation. (In the remaining time, Amit presented the linearization of the dynamics and will continue the discussion on 1 March, 2005. Some background material will be posted when we return from IPAM on February 27.)

1 Motivation: A minimization problem

The discussion here follows closely the formulation in [1, Chapter 3, pp. 113-115, pp. 144-152]. Suppose we expect $\theta(t)$ to be a linear function of t, so that we expect to write the measurements $y(t)$ of θ in the form,

$$y(t) = \theta(t) + n(t) = a + bt + n(t)\,, \tag{1}$$

where $n(t)$ is "measurement noise." Given M measurements at times $t_1, t_2, \ldots, t_M$, we write M equations

$$\mathbf{Ex} + \mathbf{n} = \mathbf{y}, \tag{2}$$

where

$$\mathbf{E} = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \\ 1 & t_M \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} a \\ b \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y(t_1) \\ y(t_2) \\ \vdots \\ y(t_M) \end{bmatrix}, \quad \mathbf{n} = \begin{bmatrix} n(t_1) \\ n(t_2) \\ \vdots \\ n(t_M) \end{bmatrix}. \tag{3}$$

The "best" solution is obtained by minimizing, wrt x_i (here, $x_1 = a$ and $x_2 = b$), the quantity

$$J := \mathbf{n}^T\mathbf{n} = (\mathbf{Ex} - \mathbf{y})^T(\mathbf{Ex} - \mathbf{y})\,. \tag{4}$$

Differentiating wrt x_i leads to the *normal equations*,

$$\mathbf{E}^T\mathbf{Ex} = \mathbf{E}^T\mathbf{y}\,, \tag{5}$$

The solution is *not* $\mathbf{x} = \mathbf{E}^{-1}\mathbf{y}$ – the inverse $\mathbf{E}$ or $\mathbf{E}^T$ is not even defined for $M \neq 2$. It could be

$$\tilde{\mathbf{x}} = (\mathbf{E}^T\mathbf{E})^{-1}\mathbf{E}^T\mathbf{y}\,. \tag{6}$$

if $(\mathbf{E}^T\mathbf{E})^{-1}$ can be found under certain assumptions. An alternate approach, to be described next, is to use the *singular value decomposition* of the matrix $\mathbf{E}$. This latter approach can be beneficial when $M \neq N$, when $(\mathbf{E}^T\mathbf{E})^{-1}$ does not exist, or when the eigenvectors of $M \times M$ matrix $\mathbf{E}$ are not orthonormal. It is also helpful not only because singular values, unlike eigenvalues, exist for every matrix but also because the singular vectors tell us something about the "dynamics" under $\mathbf{E}$. (The last sentence couldn't have been more vague, but I will leave it at that...) See [1] or email (yes, that is a verb too) Juan for an extended discussion.

2 Singular values and vectors

For any $M \times N$ matrix $\mathbf{E}$, if we construct a matrix

$$\mathbf{B} = \begin{bmatrix} \mathbf{0} & \mathbf{E}^T \\ \mathbf{E} & \mathbf{0} \end{bmatrix}, \tag{7}$$

then the eigenvector equations $\mathbf{B}\mathbf{q_i} = \lambda_i \mathbf{q}_i$ can be written as

$$\mathbf{E}\mathbf{v_i} = \lambda_i \mathbf{u}_i\,, \quad \text{and} \quad \mathbf{E}^T\mathbf{u_i} = \lambda_i \mathbf{v}_i\,, \quad \text{where} \quad \mathbf{q}_i = \begin{bmatrix} \mathbf{v_i} \\ \mathbf{u_i} \end{bmatrix}, \tag{8}$$

which are equivalent to $\mathbf{E}^T\mathbf{E}\mathbf{v}_i = \lambda_i^2\mathbf{v}_i$ and $\mathbf{E}\mathbf{E}^T\mathbf{u}_i = \lambda_i^2\mathbf{u}_i$. If $\mathbf{U}$ and $\mathbf{V}$ are, respectively, the symmetric $M \times M$ and $N \times N$ matrices with columns $\mathbf{u}_i$ and $\mathbf{v}_i$ and $\mathbf{\Lambda}$ is the diagonal matrix with diagonal elements λ_i (with zeros filling up the necessary gaps), then it follows that

$$\mathbf{E} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T\,. \tag{9}$$

All this is relevant to the minimization problem because in the general case, the "optimal" solution to that problem can be written as

$$\tilde{\mathbf{x}} = \mathbf{V}'\mathbf{\Lambda}'^{-1}\mathbf{U}'\mathbf{y}\,, \tag{10}$$

where the $'$ indicates that we have taken only the parts of these matrices that correspond to the non-zero λ_i values.

The vectors $\mathbf{u}_i$ and $\mathbf{v}_i$ are called (backward and forward, see [2]) *singular vectors* while λ_i are called the *singular values* of $\mathbf{E}$.

The relevance of all this to data assimilation is that the singular values and vectors of the tangent linear model provide one of the methods for generating ensembles. It can be shown that these vectors sample the "relevant" directions in the phase space and the proof (either mathematical, numerical, or simply by assertion) is left as an exercise to the reader. We will fearlessly move on to the next topic which has something to do with Lyapunov vectors, bred vectors and their relation to each other and to singular vectors.

(In the remaining time, Amit presented the linearization of the dynamics and will continue the discussion on 1 March, 2005. Some background material will be posted when we return from IPAM on February 27.)

References

[1] C. Wunsch, *The Ocean Circulation Inverse Problem*, Cambridge University Press (1996).

[2] B. Legras and R. Vautard, "A guide to Liapunov vectors," in *ECMWF seminar proceedings: Predictability, Vol I* 4-8 September 1995 `http://www.samsi.info/200405/data/activity/groups/lowdimensional/papers/LegrasV96.pdf` (Scanned local copy)