Environmental Modelling & Software 37 (2012) 41–54

Contents lists available at SciVerse ScienceDirect

Environmental Modelling & Software

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

Using tensor calculus for scenario modelling

Manfred Lenzen [a,1], Bonnie McBain [b,*]

[a] ISA, School of Physics A28, The University of Sydney, NSW 2006, Australia
[b] School of Environmental and Life Sciences, Discipline of Geography, Bldg SRR, University of Newcastle, Callaghan NSW 2308, Australia

ARTICLE INFO

Article history:
Received 22 February 2011
Received in revised form
23 February 2012
Accepted 25 February 2012
Available online 7 April 2012

Keywords:
Scenario modelling
Finite differences
Sparse tensors
Tensor calculus

ABSTRACT

We have demonstrated the benefits of sparse tensor calculus for finite-difference techniques that are widely applied to Integrated Assessment (IA). Using a tensor toolbox for Matlab, we have developed efficient code for progressing a system of state variables connected by a large variety of interaction types. Using a small example of twenty variables across three countries, we demonstrate how the tensor formalism allows not only for compact and fast scenario modelling, but also for straightforward implementation of sensitivity and Monte-Carlo analyses, as well as Structural Decomposition Analysis. In particular, we show how sparse tensor code can be exploited in order to search for potentially important, but yet unknown relationships in the interaction network between all variables.

1. Introduction

There is a wealth of approaches to modelling the interactions between human activities and the environment. Given that such models transcend the traditional disciplines, scientific activity has over the past years banded together under the label Integrated Assessment (IA). There now exists a number of IA programs around the world, large-scale with respect to the model itself, but also with respect to the size of team and budget. Dowlatabadi (1995) presents a typology of approaches, listing techniques such as general equilibrium, Linear Programming, probabilistic models, and decision trees. In many cases, IA models are an assembly of coupled sub-modules that each look after a particular aspect of the entire model (for instance see Schaldach et al., 2011). Schaldach and Priess (2008) provide an overview of integrated land system models.

This article is concerned with evaluating the performance and advantages of sparse tensor algebra for Integrated Assessment models of the kind that (Ha-Duong, 1997) describes as "pushed by the past". (Ha-Duong, 1997) (p. 2) defines these as "*[models] that treated in the framework of differential equations or finite difference equations, [...]* where the flow of the calculation follows the natural time: Given the state of the system at date 0, state at date 1 is computed first, then date 2 is examined, recursively up to date T". In IA, these finite-difference techniques are often used for the climate and air pollution modules of large coupled models (see for example (McPherson et al., 2003) and (Valverde, 2005)). In this work, we evaluate the mechanism of pushing, or progressing system state variables over time, using a finite-difference formulation controlled by a multi-order sparse progression tensor. This tensor represents the linearization of various non-linear functions describing connecting pairs of variables into a cause-and-effect relationship.

Braddock et al. (1995) and Zapert et al. (1998) use a second order matrix expression in order to analyse stochasticity in the IMAGE model. They use a vector state variable $\mathbf{X}(t)$ incorporating 159 variables of the model, a linear interaction matrix $\mathbf{A}$, a vector function $\mathbf{N}$ encompassing non-linear relationships, and a constant-coefficient term $\mathbf{U}$ containing the exogenous forcing terms. Our work follows this general set-up. Also similar to this work, these authors linearise the non-linear terms by constructing the gradient matrix $\nabla\boldsymbol{N}$. The main aspect in (Braddock et al., 1995) is that the authors subject parts of the interaction matrix $\mathbf{A}$ to random perturbations in order to examine the stochastic properties of the system.

Whilst stochasticity can also easily be evaluated using sparse tensor algebra, we do not explore this issue in this article, since (Braddock et al., 1995) already have provided a comprehensive analysis. Instead we concentrate on the following aspects: we demonstrate how sparse tensor algebra can a) lead to an extremely compact computer code, b) enable efficient sensitivity and structural decomposition analysis, and c) provide a means of

* Corresponding author. Tel.: +61 02 4921 7484; fax: +61 02 4921 5472.
E-mail addresses: m.lenzen@physics.usyd.edu.au (M. Lenzen), Bonnie.Mcbain@newcastle.edu.au (B. McBain).
[1] Tel.: +61 0 2 9351 5985.

1364-8152/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.envsoft.2012.02.020

exploring the potential effects of yet unknown relationships between variables, by introducing spurious coefficients into the tensor.

In the following, we will introduce the tensor formulation of some finite-difference schemes in Section 2, provide a small-scale and simplified, but realistic example in Section 3, and conclude in Section 4.

2. Formulating causal relationships in tensor finite-difference notation

2.1. Finite differences

In essence, the finite difference method (Rübenkönig, 2006) is a numeric method that is used to solve partial differential equations by approximating derivatives as differences, that is

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \approx \frac{f(x+\Delta x) - f(x)}{\Delta x} \quad \text{if } \Delta x \text{ is "small"}. \tag{1}$$

For example, approximating the differential equation $\dot{y} = \partial y/\partial t = Cy$ leads to the finite-difference equation

$$y(t+\Delta t) = y(t) + y(t)C\Delta t. \tag{2}$$

Starting with a known $y(t_0)$ and a rate of change C, the function y can be progressed over time using Eq. (2). The well-known analytical solution to this equation is the exponential function $y(t) = y(t_0)e^{Ct}$. y changes exponentially at a given rate C.

Approximating the differential equation $\dot{y}/y = C(\dot{x}/x)$ leads to the finite-difference equation

$$y(t+\Delta t) = y(t) + y(t)C\frac{x(t+\Delta t) - x(t)}{x(t)}. \tag{3}$$

The solution to this equation is the power function $y(t) = (y(t_0)/x(t_0)^C)x(t)^C$. If x changes 1%, y changes C%. C is called the x -elasticity of y. Table 1 provides an overview of some basic differential equations and their finite-difference forms.

2.2. Extensions: relative change variants, temporal and spatial lags

There are a number of modifications and special cases of the finite-difference equations in Table 1. First, the relative change term $[x(t+\Delta t) - x(t)]/x(t)$ references the change in x between times t and $t+\Delta t$ to the state of x at time t. In analogy to this Laspeyres form, it is equally possible to use Marshall-Edgeworth and Paasche variants of indexation (Lenzen, 2006) as in Eqs. (4) and (5), respectively .

$$[x(t+\Delta t) - x(t)]/0.5[x(t+\Delta t) + x(t)] \tag{4}$$

$$[x(t+\Delta t) - x(t)]/x(t+\Delta t) \tag{5}$$

In general,

$$\frac{\Delta x}{x}(t+\Delta t) = \frac{x(t+\Delta t) - x(t)}{\alpha x(t+\Delta t) + (1-\alpha)x(t)}, \quad \text{with } 0 \le \alpha \le 1. \tag{6}$$

The same variants apply to the term $\dot{y}/y$ in Table 1. Further, $y(t)$ may depend on the state, or the change in state of x at a time $t-d$ prior to t. In other words, there may be a temporal delay, or lag, of size d in the effect of x, or a change in x, on y. Finite-difference equations can readily be modified to accommodate such lags by setting

$$\frac{\Delta x}{x}(t+\Delta t - d) = \frac{x(t+\Delta t - d) - x(t-d)}{\alpha x(t+\Delta t - d) + (1-\alpha)x(t-d)}, \quad \text{with } 0 \le \alpha \le 1. \tag{7}$$

In order for the iterations to work, the minimum delay must be $d_{\min} = \Delta t$. A well-known example for temporal lags is the causal chain between greenhouse gas emissions, atmospheric concentration, and global temperature change, where response functions describe delays of more than 100 years (Meira and Miguez, 2000).

A special case of a temporal lag is given when a system of causal variables contains loops. Assume for example a causal network $x \to y \to z$ $\nwarrow q \swarrow$ where x influences y. which in turn influences z which in turn influences q which then feeds back onto y. In the finite difference approach the progression to $y(t+\Delta t)$ can be evaluated from $x(t+\Delta t)$, but not from $q(t+\Delta t)$,, because this quantity is not yet known at the time that $x(t+\Delta t)$ is known. Such a feedback loop can only be evaluated with q lagging one time step behind y.

Finally, in spatially explicit scenario modelling, a variable may influence itself across space. In other words, the state, or change in state, of a variable $x(t,p)$ at point p may have a direct causal effect on the state, or change in state, of $x(t,p')$ at point p'. Relationships of such kind are called spatial autocorrelation, or spatial lags, and finite differences can be written for example as

$$x(t+\Delta t, p') - x(t, p') = C[x(t,p) - x(t-\Delta t, p)]. \tag{8}$$

Note that a temporal lag of one time step is once again unavoidable. An example for a spatially autocorrelated variable is the number of threatened species in a particular country, which is dependent on the number of threatened species in neighbouring countries (Pandit and Laband, 2007). The reason for this autocorrelation is that species migrate, and hence the degradation or destruction of habitat has cross-border effects.

Table 1
Overview of finite-difference formulations. Differential equations can be read by linking row trailers and column headers at the "…" sign. Each intersection provides row-wise: a) a label for the relationship, b) a finite-difference expression for $\Delta y = y(t+\Delta t) - y(t)$, c) the analytical solution for $y(t)$, and d) a description of the causal relationship determining y.

	$\ldots C$	$\ldots Ct$	$\ldots C\dot{x}$	$\ldots C\frac{\dot{x}}{x}$
$\dot{y} = \ldots$	Linear $C\Delta t$ $Ct + y(t_0)$ y changes linearly over time at a constant rate C	Quadratic $Ct\Delta t$ $Ct^2 + y(t_0)$ y changes over time at a linearly changing rate Ct	Proportional $C[x(t+\Delta t) - x(t)]$ $C[x(t) - x(t_0)] + y(t_0)$ y changes proportional to x	Logarithmic $C[x(t+\Delta t) - x(t)]/x(t)$ $C\ln[x(t) - x(t_0)] + y(t_0)$ y changes proportional to relative changes in x
$\frac{\dot{y}}{y} = \ldots$	Malthusian $y(t)C\Delta t$ $y(t_0)e^{Ct}$ y changes exponentially over time at a constant rate C	Gaussian $y(t)Ct\Delta t$ $y(t_0)\ e^{Ct^2/2}$ y changes exponentially over time at a linearly changing rate Ct	Exponential $y(t)C[x(t+\Delta t) - x(t)]$ $y(t_0)e^{C[x(t)-x(t_0)]}$ y changes relatively proportional to x	Elastic $y(t)C[x(t+\Delta t) - x(t)]/x(t)$ $y(t_0)/x(t_0)^C x(t)^C$ y changes relatively proportional to relative changes in x