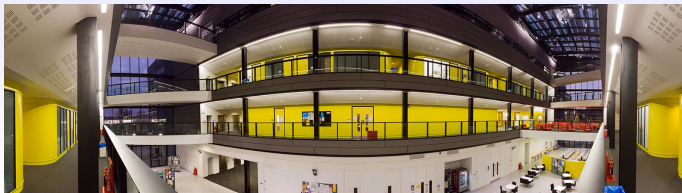


Functions of Matrices and Nearest Correlation Matrices

Nick Higham
School of Mathematics
The University of Manchester

`higham@ma.man.ac.uk`
`http://www.ma.man.ac.uk/~higham/`



What is a Matrix Function?

It's *not*

- $\det(A)$ or $\text{trace}(A)$,
- elementwise evaluation: $f(a_{ij})$,
- A^T ,
- matrix factor (e.g., $A = LU$).

What is a Matrix Function?

It's *not*

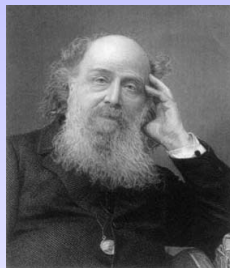
- $\det(A)$ or $\text{trace}(A)$,
- elementwise evaluation: $f(a_{ij})$,
- A^T ,
- matrix factor (e.g., $A = LU$).

It *is*

- A^{-1} ,
- e^A ,
- \sqrt{A} ,
- ...

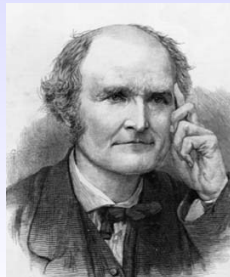
Cayley and Sylvester

- Term “**matrix**” coined in 1850 by James Joseph Sylvester, FRS (1814–1897).



- **Matrix algebra** developed by Arthur Cayley, FRS (1821–1895).

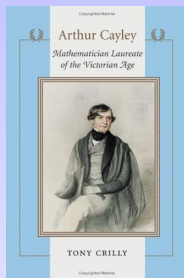
Memoir on the Theory of Matrices (1858).



Cayley and Sylvester on Matrix Functions

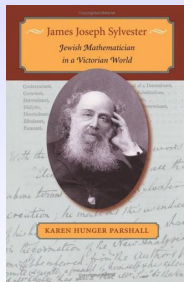
- Cayley considered matrix square roots in his 1858 memoir.

Tony Crilly, *Arthur Cayley: Mathematician Laureate of the Victorian Age*, 2006.



- Sylvester (1883) gave first definition of $f(A)$ for general f .

Karen Hunger Parshall, *James Joseph Sylvester. Jewish Mathematician in a Victorian World*, 2006.



Two Definitions

Definition (Taylor series)

If f has a Taylor series expansion $f(z) = \sum_{k=0}^{\infty} a_k z^k$ with radius of convergence r and $\rho(A) < r$ then

$$f(A) = \sum_{k=0}^{\infty} a_k A^k.$$

Two Definitions

Definition (Taylor series)

If f has a Taylor series expansion $f(z) = \sum_{k=0}^{\infty} a_k z^k$ with radius of convergence r and $\rho(A) < r$ then

$$f(A) = \sum_{k=0}^{\infty} a_k A^k.$$

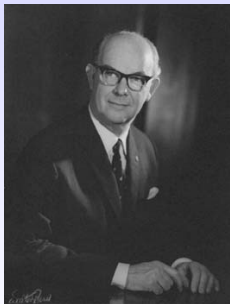
Definition (Cauchy integral formula)

$$f(A) = \frac{1}{2\pi i} \int_{\Gamma} f(z)(zI - A)^{-1} dz,$$

where f analytic on and inside closed contour Γ enclosing $\lambda(A)$.

Matrices in Applied Mathematics

- Frazer, Duncan & Collar, Aerodynamics Division of NPL: aircraft flutter, matrix structural analysis.
- **Elementary Matrices & Some Applications to Dynamics and Differential Equations, 1938.**
Emphasizes importance of e^A .
- Arthur Roderick Collar, FRS (1908–1986): *“First book to treat matrices as a branch of applied mathematics”*.



Solving Ordinary Differential Equations

$$\frac{d^2y}{dt^2} + Ay = 0, \quad y(0) = y_0, \quad y'(0) = y'_0$$

has solution

$$y(t) = \cos(\sqrt{A}t)y_0 + (\sqrt{A})^{-1} \sin(\sqrt{A}t)y'_0.$$

Solving Ordinary Differential Equations

$$\frac{d^2y}{dt^2} + Ay = 0, \quad y(0) = y_0, \quad y'(0) = y'_0$$

has solution

$$y(t) = \cos(\sqrt{A}t)y_0 + (\sqrt{A})^{-1} \sin(\sqrt{A}t)y'_0.$$

But also

$$\begin{bmatrix} y' \\ y \end{bmatrix} = \exp \left(\begin{bmatrix} 0 & -tA \\ tI_n & 0 \end{bmatrix} \right) \begin{bmatrix} y'_0 \\ y_0 \end{bmatrix}.$$

Phi Functions: Definition

$$\varphi_0(z) = e^z, \quad \varphi_1(z) = \frac{e^z - 1}{z}, \quad \varphi_2(z) = \frac{e^z - 1 - z}{z^2}, \dots$$

$$\varphi_{k+1}(z) = \frac{\varphi_k(z) - 1/k!}{z}.$$

$$\varphi_k(z) = \sum_{j=0}^{\infty} \frac{z^j}{(j+k)!}.$$

Phi Functions: Solving ODEs

$$y \in \mathbb{C}^n, A \in \mathbb{C}^{n \times n}.$$

$$\frac{dy}{dt} = Ay, \quad y(0) = y_0 \quad \Rightarrow \quad y(t) = e^{At}y_0.$$

Phi Functions: Solving ODEs

$$y \in \mathbb{C}^n, A \in \mathbb{C}^{n \times n}.$$

$$\frac{dy}{dt} = Ay, \quad y(0) = y_0 \quad \Rightarrow \quad y(t) = e^{At}y_0.$$

$$\frac{dy}{dt} = Ay + b, \quad y(0) = 0 \quad \Rightarrow \quad y(t) = t\varphi_1(tA)b.$$

Phi Functions: Solving ODEs

$$y \in \mathbb{C}^n, A \in \mathbb{C}^{n \times n}.$$

$$\frac{dy}{dt} = Ay, \quad y(0) = y_0 \quad \Rightarrow \quad y(t) = e^{At}y_0.$$

$$\frac{dy}{dt} = Ay + b, \quad y(0) = 0 \quad \Rightarrow \quad y(t) = t\varphi_1(tA)b.$$

$$\frac{dy}{dt} = Ay + ct, \quad y(0) = 0 \quad \Rightarrow \quad y(t) = t^2\varphi_2(tA)c.$$

⋮

Exponential Integrators

Consider

$$y' = Ly + N(y).$$

$N(y(t)) \approx N(y(0))$ implies

$$y(t) \approx e^{tL}y_0 + t\varphi_1(tL)N(y(0)).$$

Exponential Euler method:

$$y_{n+1} = e^{hL}y_n + h\varphi_1(hL)N(y_n).$$

Lawson (1967); recent resurgence.

Toolbox of Matrix Functions

- Want software for evaluating interesting f at matrix args as well as scalar args.
- MATLAB has `expm`, `logm`, `sqrtn`, `funm`.
- **The Matrix Function Toolbox** (H, 2008).
- NAG Library:
 - `f01ecf` (`f01ecc`) for matrix exponential.
 - `f01eff`/`f01fff` for function of symmetric/Hermitian matrix.
 - More on the way ...

Scaling and Squaring Method

Scale: $B \leftarrow A/2^s$ so $\|B\|_\infty \approx 1$

Approximate: $r_m(B) = [m/m]$ Padé approximant to e^B

Square: $X = r_m(B)^{2^s} \approx e^A$

- **Moler & Van Loan (1978)** “*Nineteen dubious ways to compute the exponential of a matrix*”—methodology for choosing s and m .
- **H (2005)**: sharper analysis giving optimal s and m .
- **Al-Mohy & H (2009)**: further improvements.

Compute $e^A b$

Exploit, for integer s ,

$$e^A b = (e^{s^{-1}A})^s b = \underbrace{e^{s^{-1}A} e^{s^{-1}A} \dots e^{s^{-1}A}}_{s \text{ times}} b.$$

Choose s so $T_m(s^{-1}A) = \sum_{j=0}^m \frac{(s^{-1}A)^j}{j!} \approx e^{s^{-1}A}$. Then

$$b_{i+1} = T_m(s^{-1}A)b_i, \quad i = 0: s-1, \quad b_0 = b$$

yields $b_s \approx e^A b$.

Al-Mohy & H (2011), SIAM J. Sci. Comp.

Experiment

Compute $e^{tA}b$ for Harwell–Boeing matrices:

- **orani678**, $n = 2529$, $t = 100$, $b = [1, 1, \dots, 1]^T$;
- **bcspwr10**, $n = 5300$, $t = 10$, $b = [1, 0, \dots, 0, 1]^T$.

2D Laplacian matrix, **poisson**. $\text{tol} = 6 \times 10^{-8}$.

	Alg AH			ode15s		
	time	cost	error	time	cost	error
orani678	0.13	878	4e-8	136	7780+...	2e-6
bcspwr10	0.021	215	7e-7	2.92	1890+...	5e-5
poisson	3.76	29255	2e-6	2.48	402+...	8e-6
4 poisson	15	116849	9e-6	3.24	49+...	1e-1

General Functions

- **Schur–Parlett algorithm (Davies & H, 2003)**
computes $f(A)$ given the ability to evaluate $f^{(k)}(x)$ for any k and x .
- Implemented in MATLAB's **funm**.
- Beware *unstable diagonalization algorithm*:

```
function F = funm_ev(A, fun)
%FUNM_EV    Evaluate general matrix
%           function via eigensystem.
[V,D] = eig(A);
F = V * diag(feval(fun,diag(D))) / V;
```

Email from a Power Company

The problem has arisen through proposed methodology on which the company will incur charges for use of an electricity network.

⋮

I have the use of a computer and Microsoft Excel.

⋮

*I have an Excel spreadsheet containing the transition matrix of how a company's [Standard & Poor's] credit rating changes from one year to the next. I'd like to be working in eighths of a year, so the aim is to find the **eighth root of the matrix.***

Chronic Disease Example

- Estimated 6-month transition matrix.
- Four AIDS-free states and 1 AIDS state.
- 2077 observations (Charitos et al., 2008).

$$P = \begin{bmatrix} 0.8149 & 0.0738 & 0.0586 & 0.0407 & 0.0120 \\ 0.5622 & 0.1752 & 0.1314 & 0.1169 & 0.0143 \\ 0.3606 & 0.1860 & 0.1521 & 0.2198 & 0.0815 \\ 0.1676 & 0.0636 & 0.1444 & 0.4652 & 0.1592 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Want to estimate the **1-month transition matrix**.

$$\Lambda(P) = \{1, 0.9644, 0.4980, 0.1493, -0.0043\}.$$

- **H & Lin** (2011).
- **Lin** (2011, Chap. 3 for survey of regularization methods).

MATLAB: Arbitrary Powers

```
>> A = [1 1e-8; 0 1]
```

```
A =
```

```
1.0000e+000    1.0000e-008
              0    1.0000e+000
```

```
>> A^0.1
```

```
ans =
```

```
1    0
0    1
```

```
>> expm(0.1*logm(A))
```

```
ans =
```

```
1.0000e+000    1.0000e-009
              0    1.0000e+000
```

MATLAB Arbitrary Power

- New **Schur algorithm** (H & Lin, 2011) reliably computes A^p for any real p .
- Brings improvements over MATLAB A^p even for negative integer p .
- Alternative Newton-based algorithms available for $A^{1/q}$ with q an integer, e.g., for

$$X_{k+1} = \frac{1}{q} [(q+1)X_k - X_k^{q+1}A] \quad X_0 = A,$$

$$X_k \rightarrow A^{-1/q}.$$

- University of Manchester and NAG (2010–2013) funded by EPSRC, NAG and TSB.
- Developing suite of NAG Library codes for matrix functions.
- KTP Associate Edvin Deadman.
- Improvements to existing state of the art.
- Suggestions for prioritizing code development welcome.

- **Functions of Matrices: Theory and Computation**, 2011–2015 (value €**2M**).
- 3 postdocs, 2 PhD students, international visitors, 2 workshops.
- New algorithms will be developed.

Questions From Finance Practitioners

“Given a real symmetric matrix A which is almost a correlation matrix what is the best approximating (in Frobenius norm?) correlation matrix?”

“I am researching ways to make our company’s correlation matrix positive semi-definite.”

“Currently, I am trying to implement some real options multivariate models in a simulation framework. Therefore, I estimate correlation matrices from inconsistent data set which eventually are non psd.”

Correlation Matrix

An $n \times n$ symmetric positive semidefinite matrix A with $a_{ij} \equiv 1$.

Properties:

- symmetric,
- 1s on the diagonal,
- off-diagonal elements between -1 and 1 ,
- eigenvalues nonnegative.

Correlation Matrix

An $n \times n$ symmetric positive semidefinite matrix A with $a_{ij} \equiv 1$.

Properties:

- symmetric,
- 1s on the diagonal,
- off-diagonal elements between -1 and 1 ,
- eigenvalues nonnegative.

Is this a correlation matrix?

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Correlation Matrix

An $n \times n$ symmetric positive semidefinite matrix A with $a_{ij} \equiv 1$.

Properties:

- symmetric,
- 1s on the diagonal,
- off-diagonal elements between -1 and 1 ,
- eigenvalues nonnegative.

Is this a correlation matrix?

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

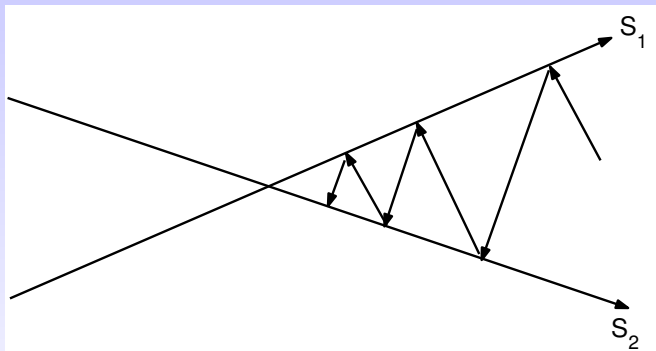
Spectrum: $-0.4142, 1.0000, 2.4142$.

How to Proceed

- ✗ Make ad hoc modifications to matrix: e.g., shift negative e'vals up to zero then diagonally scale.
- ✓ Plug the gaps in the missing data, then compute an exact correlation matrix.
- ✓ Compute the **nearest correlation matrix** in the weighted Frobenius norm ($\|A\|_F^2 = \sum_{i,j} w_i w_j a_{ij}^2$).
 - Constraint set is a closed, convex set, so unique minimizer.

Alternating Projections

von Neumann (1933), for subspaces.



Dykstra (1983) incorporated corrections for closed convex sets.

Newton Method

- **Qi & Sun** (2006): convergent Newton method based on theory of **strongly semismooth matrix functions**. **Globally** and **quadratically** convergent.
- Various algorithmic improvements by Borsdorf & H (2010).
- Implemented in NAG codes **g02aaf** (**g02aac**) and **g02abf** (weights, lower bound on ei'vals—Mark 23).

Factor Model (1)

$$\xi = \underbrace{X}_{n \times k} \underbrace{\eta}_{k \times 1} + \underbrace{F}_{n \times n} \underbrace{\varepsilon}_{n \times 1}, \quad \eta_i, \varepsilon_i \in N(0, 1),$$

where $F = \text{diag}(f_{ii})$. Implies

$$\sum_{j=1}^k x_{ij}^2 \leq 1, \quad i = 1 : n.$$

- “Multifactor normal copula model”.
- Collateralized debt obligations (CDOs).
- Multivariate time series.

Factor Model (2)

Yields correlation matrix of form

$$C(X) = D + XX^T = D + \sum_{j=1}^k x_j x_j^T,$$

$$D = \text{diag}(I - XX^T), \quad X = [x_1, \dots, x_k].$$

$C(X)$ has **k factor correlation matrix structure**.

$$C(X) = \begin{bmatrix} 1 & y_1^T y_2 & \cdots & y_1^T y_n \\ y_1^T y_2 & 1 & \cdots & \vdots \\ \vdots & \cdots & \ddots & y_{n-1}^T y_n \\ y_1^T y_n & \cdots & y_{n-1}^T y_n & 1 \end{bmatrix}, \quad y_i \in \mathbb{R}^k.$$

All your hedges in one basket

Leif Andersen, Jakob Sidenius and Susanta Basu present new techniques for single-tranche CDO sensitivity and hedge ratio calculations. Using factorisation of the copula correlation matrix, discretisation of the conditional loss distribution followed by a recursion-based probability calculation, and derivation of analytical formulas for deltas, they demonstrate a significant improvement in computational speeds

In a traditional synthetic collateralised debt obligation (CDO), the arranger tranches out credit losses on a pool of credit default swaps (CDSs) and passes them through to different investors. Assuming that investors for all tranches can be identified, the arranger is typically left with fairly moderate market exposure. For various reasons, placing the entire pool capital structure with investors has become increasingly difficult, and many recent credit basket derivatives expose the dealer to significant market risk. For instance, the recent 'single-tranche' CDO (STCDO) product involves the sale of a single CDO tranche to a single customer, leaving it to the arranger to manage the risk of the remaining capital structure. As STCDOs and similar 'custom' products offer significant customer benefits and are much less difficult to originate than traditional CDOs, such products are likely to increase in importance. This is especially true for managed trades where the customer has certain rights to alter the composition of the reference portfolio over time.

A basic prerequisite for active management of the risk of a credit basket derivative is the ability to accurately calculate the sensitivity of the security with respect to market and model parameters, most prominently the par CDS spreads of the underlying reference pool. The numbers of such sensitivities can be very large – many thousands – and can put considerable strain on computing resources. Moreover, the calculation of each of

where Q is the risk-neutral probability measure and λ_k is a (forward) default hazard rate function. The functions $p_k(T)$, $k = 1, \dots, N$ can be bootstrapped by standard means from the quoted CDS spreads and are assumed known for all T .

Equation (1) fully establishes the risk-neutral marginal distribution of each default time τ_k . To construct the joint distribution of all default times, we here choose² to employ a Student- t copula, which we quickly define for reference. Defining vectors $\boldsymbol{\tau} = (\tau_1, \dots, \tau_N)^T$ and $\mathbf{T} = (T_1, \dots, T_N)^T$, the joint default time distribution in the Student- t copula, becomes:

$$Q(\boldsymbol{\tau} \leq \mathbf{T}) = t_{N,v}(\tau_1^{-1}(p_1(T_1)), \dots, \tau_N^{-1}(p_N(T_N))) \quad (2)$$

where $t_{1,v}$ and $t_{N,v}$ are the one- and N -dimensional cumulative Student- t distribution functions with v degrees of freedom, respectively. Recall that the density $\eta_{N,v}$ of an N -dimensional Student- t distribution with correlation matrix Σ is:

$$\eta_{N,v}(\mathbf{z}) = C_{N,v} \left(1 + \mathbf{v}^{-1} \mathbf{z}^T \Sigma^{-1} \mathbf{z}\right)^{-\frac{v+N}{2}}, \quad C_{N,v} = \frac{\Gamma\left(\frac{v+N}{2}\right)}{\Gamma\left(\frac{v}{2}\right) \sqrt{|\Sigma|} (v\pi)^N} \quad (3)$$

where Γ is the gamma function. For high degrees of freedom, (3) approaches

k Factor Problem

$$\min_{X \in \mathbb{R}^{n \times k}} f(x) := \|A - C(X)\|_F^2 \quad \text{subject to} \quad \sum_{j=1}^k x_{ij}^2 \leq 1.$$

- Nonlinear objective function with convex quadratic constraints.
- Some existing algs ignore the constraints.

- Algorithm based on **spectral projected gradient method** (Borsdorf, H & Raydan, 2011).
 - Respects the constraints, exploits their convexity, and converges to a feasible stationary point.
 - NAG routine **g02aef** (Mark 23).
- **Principal factors method** (Andersen et al., 2003) has no convergence theory and can converge to an incorrect answer.




Conclusions

- Matrix functions a powerful and versatile tool, with excellent algs available.
- Beware unstable/impractical algs in literature!




- MATLAB $f(A)$ algs were last updated 2006.
- Currently implementing state of the art $f(A)$ algs for NAG Library via the KTP.

- Excellent algs available for nearest correlation matrix problems.
- Beware algs in literature that may not converge or converge to wrong solution!



References I

-  A. H. Al-Mohy and N. J. Higham.
A new scaling and squaring algorithm for the matrix exponential.
SIAM J. Matrix Anal. Appl., 31(3):970–989, 2009.
-  A. H. Al-Mohy and N. J. Higham.
Computing the action of the matrix exponential, with an application to exponential integrators.
SIAM J. Sci. Comput., 33(2):488–511, 2011.
-  L. Anderson, J. Sidenius, and S. Basu.
All your hedges in one basket.
Risk, pages 67–72, Nov. 2003.
| www.risk.net |.

References II

-  R. Borsdorf and N. J. Higham.
A preconditioned Newton algorithm for the nearest correlation matrix.
IMA J. Numer. Anal., 30(1):94–107, 2010.
-  R. Borsdorf, N. J. Higham, and M. Raydan.
Computing a nearest correlation matrix with factor structure.
SIAM J. Matrix Anal. Appl., 31(5):2603–2622, 2010.
-  T. Charitos, P. R. de Waal, and L. C. van der Gaag.
Computing short-interval transition matrices of a discrete-time Markov chain from partially observed data.
Statistics in Medicine, 27:905–921, 2008.

References III

-  T. Crilly.
Arthur Cayley: Mathematician Laureate of the Victorian Age.
Johns Hopkins University Press, Baltimore, MD, USA,
2006.
ISBN 0-8018-8011-4.
xxi+610 pp.
-  P. I. Davies and N. J. Higham.
A Schur–Parlett algorithm for computing matrix
functions.
SIAM J. Matrix Anal. Appl., 25(2):464–485, 2003.

References IV



R. A. Frazer, W. J. Duncan, and A. R. Collar.
Elementary Matrices and Some Applications to Dynamics and Differential Equations.
Cambridge University Press, Cambridge, UK, 1938.
xviii+416 pp.
1963 printing.



P. Glasserman and S. Suchintabandit.
Correlation expansions for CDO pricing.
Journal of Banking & Finance, 31:1375–1398, 2007.



N. J. Higham.
The Matrix Function Toolbox.
[http://www.ma.man.ac.uk/~higham/mftoolbox.](http://www.ma.man.ac.uk/~higham/mftoolbox)

References V



N. J. Higham.

The scaling and squaring method for the matrix exponential revisited.

SIAM J. Matrix Anal. Appl., 26(4):1179–1193, 2005.



N. J. Higham.

Functions of Matrices: Theory and Computation.

Society for Industrial and Applied Mathematics,
Philadelphia, PA, USA, 2008.

ISBN 978-0-898716-46-7.

xx+425 pp.

References VI



N. J. Higham.

The scaling and squaring method for the matrix exponential revisited.

SIAM Rev., 51(4):747–764, 2009.



N. J. Higham and A. H. Al-Mohy.

Computing matrix functions.

Acta Numerica, 19:159–208, 2010.

References VII



N. J. Higham and L. Lin.

A Schur–Padé algorithm for fractional powers of a matrix.

MIMS EPrint 2010.91, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, Oct. 2010.

25 pp.

To appear in *SIAM J. Matrix Anal. Appl.*





N. J. Higham and L. Lin.

On p th roots of stochastic matrices.

Linear Algebra Appl., 435(3):448–463, 2011.

References VIII

-  J. D. Lawson.
Generalized Runge-Kutta processes for stable systems with large Lipschitz constants.
SIAM J. Numer. Anal., 4(3):372–380, Sept. 1967.
-  L. Lin.
Roots of Stochastic Matrices and Fractional Matrix Powers.
PhD thesis, The University of Manchester, Manchester, UK, 2010.
117 pp.
MIMS EPrint 2011.9, Manchester Institute for Mathematical Sciences.

References IX



K. H. Parshall.

James Joseph Sylvester. Jewish Mathematician in a Victorian World.

Johns Hopkins University Press, Baltimore, MD, USA,
2006.

ISBN 0-8018-8291-5.

xiii+461 pp.