ISSN 0973-9777

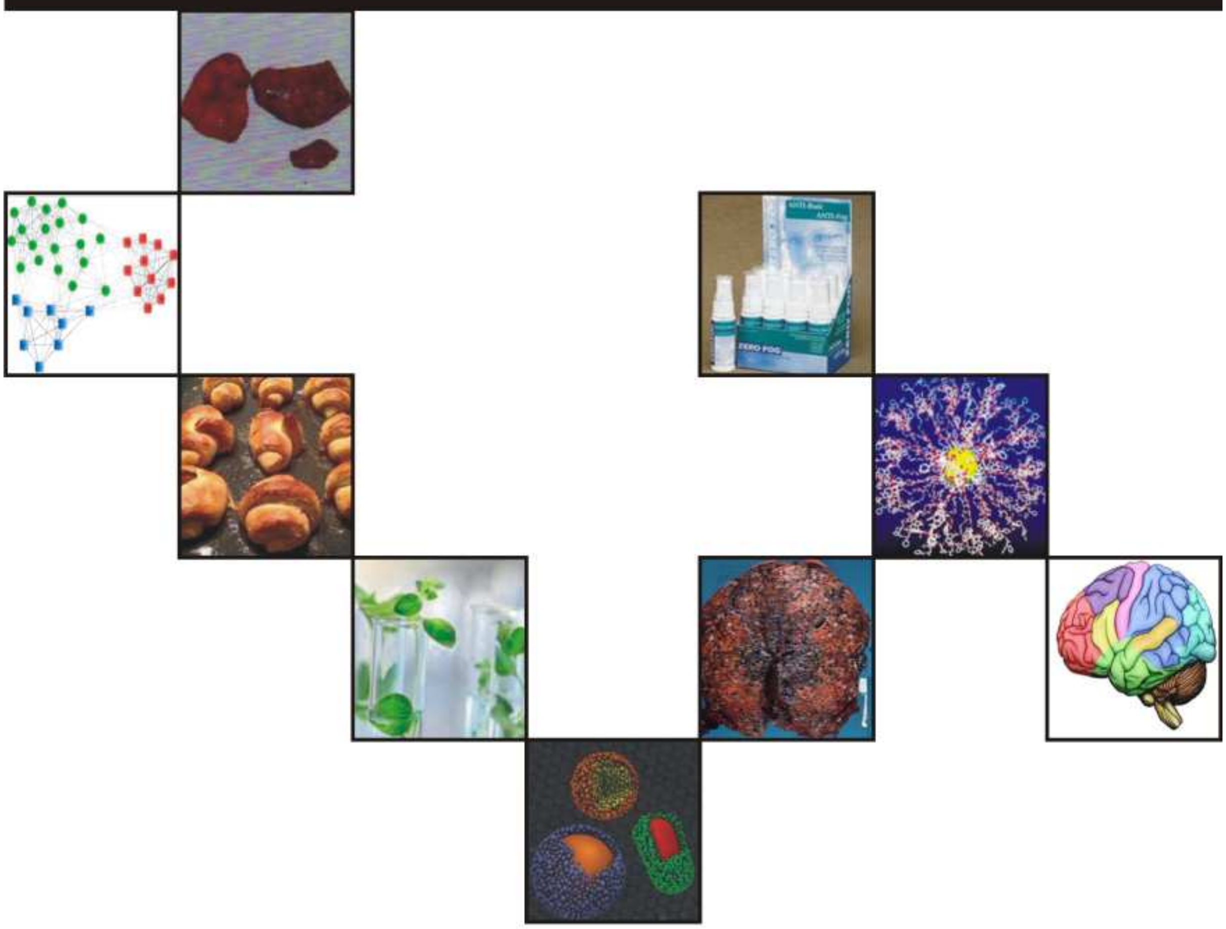
Volume-6 Number-3 May-June 2012

The Indian Journal of Research

ANVIKSHIKI

Bi-monthly International Journal of all Research

Science



MPASVO

Published on behalf of the MPASVO in association with the Member's of Anvikshiki

MANEESHA PUBLICATIONS
www.anvikshikijournal.com

Anvikshiki
The Indian Journal of Research

Bi-Monthly International Journal of All Research

Abstracts and Indexing

http//nkrc.niscair.res.in/browseByTitle.php?Keword=A, ICMJE ICMJE ,www.icmje.org, Academia.edu, banaras.academia.edu, ebookbrowse.com, BitLibrary! http:/ www.bitlib.net/, Tech eBooks ,freetechebooks.com, artapp.net, artapp.net,Catechu PDF / printfu.org, File Away,www.fileaway.info, docslibrary,http://www.docslibrary.com, ,Android Tips, Apps, Theme and Phone Reviews http://dandroidtips.com, , http://www.edu-doc.com, www.themarketingcorp.com,Dunia Ebook Gratis duniaebook.net, ,www.cn.doc-cafes.com., Google ,http://scholar.google.co.in, Website : www.onlineijra.com.Motilal Banarasi Das Index,Varanasi, Motilal Banarasi Das Index,Delhi. Banaras Hindu University Journal Index,Varanasi. www.bhu.ac.in, D.K.Publication Index, Delhi. National Institute of Science Communication and Information Resources Index, New Delhi.

Subscriptions

Anvikshiki,The Indian Journal of Research is Published every two months (January,March,May,July,September and November) by mpasvo Press,Varanasi.u.p.India. A Subscription to The Indian Journal of Research : Anvikshiki Comprises 6 Issues in Hindi and 6 in English and 3 Extra Issues. Prices include Postage by Surface mail,or For Subscription in the India by Speed Post. Airmail rates are also available on request. Annual Subscriptions Rates (Volume 3,6 Issues in Hindi,6 Issues in English and 6 Issues of science 2012):

Subscribers

Institutional : Inland 4,000 +500 Rs. P.C., Single 500+51 Rs.P.C.,Overseas 6000+2000Rs. P.C., Single 1000+500 Rs.P. C.

Personal : 2,500+500 Rs. P.C., Single 500+51 Rs. P.C., Overseas 5000+2000Rs.P.C., Single 1000+500Rs. P.C.

All correspondence related to the Journal should be addressed to

B.32/16 A., Flat No.2/1,Gopalkunj,Nariya,Lanka, Varanasi, U.P.,India

Mobile : 09935784387,Tel.0542-2310539.,e-mail : maneeshashukla76@rediffmail.com,www.anvikshikijournal.com

Office Time : 3-5 P.M.(Sunday off)

Journal set by

Maheshwar Shukla,maheshwar.shukla@rediffmail.com

9415614090

Printed by

mpasvo Press

Maneesha Publication

(Letter No.V-34564,Reg.533/2007-2008)

B-32/16-A-2/1,Gopalkunj,Nariya,Lanka

Varanasi,U.P.,India

Anvikshiki

The Indian Journal of Research

Volume 6 Number 3 May 2012

Science

Papers

PRINT ISSN 0973-9777,WEBSITE ISSN 0973-9777

Letter No.V-34564,Reg.533/2007-2008 INDIAN JOURNAL OF RESEARCH(2012)**6,106-114**
ANVIKSHIKI ISSN 0973-9777 Advance Access publication 15 Apr. 2012

MATRIX REDUCTION IN NUMERICAL OPTIMIZATION

MOHAMMAD NAVAZ RASOULIZADEH*

Declaration

The Declaration of the author for publication of Research Paper in The Indian Journal of Research Anvikshiki ISSN 0973-9777 Bi-monthly International Journal of all Research: I, *Mohammad Navaz Rasoulizadeh* the author of the research paper entitled MATRIX REDUCTION IN NUMERICAL OPTIMIZATION declare that , I take the responsibility of the content and material of my paper as I myself have written it and also have read the manuscript of my paper carefully. Also, I hereby give my consent to publish my paper in Anvikshiki journal , This research paper is my original work and no part of it or it's similar version is published or has been sent for publication anywhere else. I authorise the Editorial Board of the Journal to modify and edit the manuscript. I also give my consent to the Editor of Anvikshiki Journal to own the copyright of my research paper.

This paper develops the use of matrix reduction techniques to simplify and stabilize the solutions to various optimization problems.

1.1 Matrix Reduction

Matrix reduction approximates a matrix by removing some terms in its decomposition. Suppose that a matrix $\boldsymbol{M}$ can be expressed as asummation of matrices $\boldsymbol{M}_i$ as

$$\boldsymbol{M}=\sum_{i=1}^{k}\boldsymbol{M}_i=\boldsymbol{M}_1+\cdots+\boldsymbol{M}_k.$$

This kind of expansion i s common in matrix computation. For instance, any matrix $\boldsymbol{A}\in\mathbb{R}^{m\times n}$ has singular value decomposition(SVD).

$$\boldsymbol{A}=\boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T,$$

* Assitant Professor, Department of Mathematics, Valayat University of Iran shahr (Iran). e-Mail : navaz81@yahoo.com

Where $\boldsymbol{U}\in\mathbb{R}^{m\times m}$ and $\boldsymbol{V}\in\mathbb{R}^{n\times n}$ are orthogonal matrices, and $\boldsymbol{S}\in\mathbb{R}^{m\times n}$ is adiagonal matrix. We can write this decomposition assumption of rank one matrices

$$\boldsymbol{A}=\sum_{i=1}^{\min(m,n)} s_i \boldsymbol{u}_i \boldsymbol{v}_i^T,$$

wher e $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ are the i-th columns of $\boldsymbol{U}$ and $\boldsymbol{V}$, and s_i is the i-th diagonal element of $\boldsymbol{S}$.

If a matrix $\boldsymbol{B}\in\mathbb{R}^{m\times m}$ is symmetri can dpositive definite, Cholesky decomposition

Also generates such an expansion as

$$\boldsymbol{B}=\boldsymbol{L}\boldsymbol{L}^T=\sum_{i=1}^{m} \boldsymbol{l}_i \boldsymbol{l}_i^T,$$

where $\boldsymbol{L}$ is a lower triangular matrix, and $\boldsymbol{l}_i$ is the i-th column of $\boldsymbol{L}$. This expansion is particularly important when $\boldsymbol{B}$ is updated by alow-rank correction since this can be accomplished by adding a small number of terms to the expression.

Broadly speaking, there are two different approaches to matrix reduction. First, if

We know that only the first $\hat{k}$ matrices $\boldsymbol{M}_i$ are important to us, we can constructa reduced

Matrix $\boldsymbol{M}$ as:

$$\hat{\boldsymbol{M}}=\sum_{i=1}^{k} \boldsymbol{M}_i=\boldsymbol{M}_1+\cdots+\boldsymbol{M}_k.$$

This reduction metho discalled *truncation*.Alternatively, we can apply afiltering factor $\varphi_i\in[0,1]$to each matrix $\boldsymbol{M}_i$ for $i=1,\ldots,k$.Then,there duced matrix $\hat{\boldsymbol{M}}$ becomes

$$\hat{\boldsymbol{M}}=\sum_{i=1}^{k} \varphi_i \boldsymbol{M}_i=\varphi_1 \boldsymbol{M}_1+\cdots+\varphi_k \boldsymbol{M}_k.$$

This filtering-based reduction can be regarded as a generalized version of *truncation* since *truncation* is a special case with $\varphi_i\in\{0,1\}$.Both reduction methods are used in many applications such as regularization of ill-posed problems and factor analysis.

We can also classify the matrix reduction approaches as *build-down* and *build-up*, depending on whether were moves one terms in a given matrix expansion or we construct the reduced matrix by adding terms until acertain goal is achieved. For example, while

A complete matrix $\boldsymbol{M}$ is given to us in the problems of regression and factor analysis,we

Construct $\boldsymbol{M}$ By adding matrices $\boldsymbol{M}_i$ in constrained convex optimization.

The purposes of the matrix reduction are very different depending on particular problems. First,in regression problems in statistics,the trun cated or filtered terms are considered to benoiseor observation error, so matrix reduction purifies the given raw data.

This can be use fulinsolving least squares problems for an over-determined linear system or regularizing the solutiontoan ill-posed problem. Second,infact or analysis and principal component analysis(PCA), there duced parts are regarded as idiosyncratic (unsystematic)factors,which are not shared by multiple variables in common. Third, in constrained convex optimization problems, there duced terms might correspond to unnecessary(inactive)constraints, which do not make significant contributions to the search for an optimal solution. So,we expect a benefit of decreased computational cost by using matrix reduction.

When ever matrix reduction is applied, it is a very critical but difficult issue to decide how much to reduce the matrix. This important decision determines both the quality of there duced matrix and that of the final result. If were duce too much, we may fail to solve the problem. On the other hand,if we reduce too little,we can not expect enough benefit from there duction. It is a difficult decision because criteria for there duction must be tailored to the problem and the circumstances.For example, in regularization of ill-

$\sigma_{max}(\boldsymbol{X})$	The largest singular value of $\boldsymbol{X}$
$\sigma_{min}(\boldsymbol{X})$	The smallestsingularvalueof$\boldsymbol{X}$
$\sigma_i(\boldsymbol{X})$	The$i$-thlargestsingularvalueof$\boldsymbol{X}$
$\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x}^T\boldsymbol{x}}$	2-normforavector$\boldsymbol{x}$
$\|\boldsymbol{X}\|_2 = \sigma_{max}(\boldsymbol{X})$	2-normforamatrix$\boldsymbol{X}$
$\|\boldsymbol{X}\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}^2}$	Frobeniusnormfora matrix$\boldsymbol{X}\in R^{m\times n}$
$\mathrm{tr}(\boldsymbol{X}) = \sum_{i=1}^{n} x_{ii}$	Traceofmatrix$\boldsymbol{X}\in R^{n\times n}$
$\boldsymbol{I}_p$	Anidentitymatrixofdimension$p$

Posed problems,the criteria may change based on which distribution the embed ded noise follows,or how then oise indifferent variables is correlated.Because of this difficulty, the criteria for constraint reduction have been studied in a variety of applications.

In this dissertation,we discuss matrix reduction in three numerical optimization problems.Our study focuses on how we can determine appropriate reduction intensity for successful matrix reduction in these problems.We introduce the problems in the next section.

In addition, a few basic statistical definitions are frequently used.When a continuous random variable x has aprobability density function $p_x(x)$, the expected value$E(x)$is defined as

$$E(x) = \int_{-\infty}^{\infty} x p_x(x).$$

Then,the varianc eva r$(x)$and thest and arddeviationstd$(x)$are defined as

$$\mathrm{var}(x) = E\left((x-E(x))^2\right) = E(x^2)-(E(x))^2,$$

$$\mathrm{std}(x) = \sqrt{\mathrm{var}(x)}.$$

For two random variables x and y,the co variance co v$(x,y)$and the correlation co rr(x,y)

Are defined as

$$\text{cov}(x,y) = E((x-E(x))(y-E(y))) = E(xy)-E(x)E(y),$$

$$\text{corr}(x,y) = \frac{\text{cov}(x,y)}{\text{std}(x)\text{std}(y)}.$$

1.2 Overview of Numerical Optimization Problems
1.2.1 Total Least Squares Problems

Suppose that we have an under lying linear model,

$$(\boldsymbol{A}-\boldsymbol{E}_A)\boldsymbol{X}=(\boldsymbol{B}-\boldsymbol{E}_B),$$

Where $\boldsymbol{E}_A$ and $\boldsymbol{E}_B$ are unknown; they result from noise in the observed matrices $\boldsymbol{A}\in R^{m\times n}$ And $\boldsymbol{B}\in R^{m\times d}$. To estimate the parameters $\boldsymbol{X}$, we construct a minimization problem

$$\min_{\boldsymbol{X},\Delta\boldsymbol{A},\Delta\boldsymbol{B}} [\Delta\boldsymbol{A},\Delta\boldsymbol{B}]_F,$$

Subject to

$$(\boldsymbol{A}-\Delta\boldsymbol{A})\boldsymbol{X}=(\boldsymbol{B}-\Delta\boldsymbol{B}),$$

$$\text{Rank }([(\boldsymbol{A}-\Delta\boldsymbol{A}),(\boldsymbol{B}-\Delta\boldsymbol{B})])=r,$$

Where r is the known rank of the noise-free data$(\boldsymbol{A}-\boldsymbol{E}_A)$.

The minimization problem above can be solved by matrix reduction on the SVD of $[\boldsymbol{A},\boldsymbol{B}]$.If there we re no noise in $\boldsymbol{A}$ and $\boldsymbol{B}$,the concatenated matrix$[\boldsymbol{A},\boldsymbol{B}]$would also have

Rank r since Range $(\boldsymbol{B})\subseteq$Range$(\boldsymbol{A})$. If the rank r of the noise-free data $(\boldsymbol{A}-\boldsymbol{E}_A)$ is

Given to us,we can truncateall butther r largest singular values of $[\boldsymbol{A},\boldsymbol{B}]$.By the Eckart- Young-Mirsky Theorem , the resulting $(\boldsymbol{X},\Delta\boldsymbol{A},\Delta\boldsymbol{B})$ is the solution to the minimization problem. In addition, if the noise matrices $\boldsymbol{E}_A$ and $\boldsymbol{E}_B$ are mutually uncorrelated and have zero mean and identical standard deviations, it is known that the minimization problem

Above gives us a consistent estimate $\boldsymbol{X}$ for th eunder lying linear model.

Our study starts from the question of how we can estimate $\boldsymbol{X}$ if we do not know the rank r or if the embed ded noise matrices $\boldsymbol{E}_A$ and $\boldsymbol{E}_B$ do not have identical standard deviations and the standard deviations are unknown. If the rank$r$ is not given to us, we need to decide how many singular values to truncate. If the standard deviations of the

Noise are different and we do not know their values, we also need to find an appropriate weight α so that weighted data $\alpha\boldsymbol{A}$ and $(1-\alpha)$ $\boldsymbol{B}$ contain noise with identical standard deviations.

In this paper the author propose a method to estimate the rank r and the weight α.We also present experimental results to evaluate the proposed method.

1.2.2 Covariance Matrix Estimation

In financial portfolio theory, Mark owitz proposed the Mean-Variance(MV) portfolio problem to find an optimal portfolio of N stocks satisfying given constraints.The MV portfolio problem requires an estimated covariance matrix $\Sigma\in R^{N\times N}$for the N stock

returns. It is well known that the performance of the portfolio is very sensitive to the quality of the covariance matrix estimate,but a conventional sample covariance matrix is far from a good estimate.

The main difficulty is that the observed stock return data contain too much noise. Matrix reduction can be used to reduce the error in the covariance matrix estimate .

Suppose that we have stock return data $\boldsymbol{R} \in \mathrm{R}^{N \times T}$ of N stocks for T time periods. For

Appropriate principal component analysis (PCA), we normalize each stock return, so that larger turn values for a few stocks do not over whelm the other return values. Let $\boldsymbol{Z}$ de- note the normalized data with zero-means and identical standard deviations. From the

Singular value decomposition of $\boldsymbol{Z}$,we have

$$\boldsymbol{Z}=\boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{T}=\boldsymbol{U}\boldsymbol{F}=\sum_{i=1}^{T}\boldsymbol{u}_i\boldsymbol{f}_i^{T},$$

where $\boldsymbol{F}=\boldsymbol{S}\boldsymbol{V}^{T}$,$\boldsymbol{u}_i$ is the i- th column of $\boldsymbol{U}$, and $\boldsymbol{f}_i^{T}$ is the i-th row of $\boldsymbol{F}$. In PCA, the vector $\boldsymbol{f}_i$ is called the i-th principal component affecting the stock returns, and the vector $\boldsymbol{u}_i$ Is called a load which determines how much each stock return is affected by the i-th component. Previously, many people proposed truncating a few smallest singular values,

expecting that the principal components corresponding to the smallest singular values are more significantly contaminated by noise. However, no one has given a clear answer as to how many principal components should be truncated. This is a very difficult decision because we fundamentally do not know how many factors govern the stock returns.

In Chapter 3,we applya Tikhonov filtering function to the principal components, a monotonically increasing function of the singular value. With this smooth filtering, we expect that the influence of important principal components is amplified while potential information in less important principal components Is still preserved. Further more, we

Propose a method to determine filtering intensity. Experiments using stock return data in NYSE,AMEX, and NASDA Q from 1958 to 2007, show that the MV portfolio using Tikhonov filtered co variance matrix performs quite well.

1.2.3 Interior Point Method for Semi definite Programming

The constrained convex optimization problem known as semi definite programming (SDP) Has the following primal and dual problems:

$$\text{Primal SDP: } \min_{X} \; \boldsymbol{C} \bullet \boldsymbol{X} \quad \text{s.t. } \boldsymbol{A}_i \bullet \boldsymbol{X} = b_i \text{ for } i = 1, \ldots, m, \; \boldsymbol{X} \succeq 0,$$

$$\text{Dual SDP: } \max_{y} \; \boldsymbol{b}^{T} \boldsymbol{y} \quad \text{s.t. } \sum_{i=1}^{m} y_i \boldsymbol{A}_i + \boldsymbol{Z} = \boldsymbol{C}, \; \boldsymbol{Z} \succeq 0,$$

where $\boldsymbol{C}, \boldsymbol{A}_i, \boldsymbol{X}$, and $\boldsymbol{Z}$ are $n \times n$ symmetric matrices, $\boldsymbol{C} \bullet \boldsymbol{X} = \text{tr}(\boldsymbol{CX})$ is the trace of the matrix, and $\boldsymbol{Z} \succeq 0$ means that $\boldsymbol{Z}$ is positive semi definite.

In an interior point method (IPM) for solving the SDP, we use Newton's method to find a direction $(\Delta \boldsymbol{X}, \Delta \boldsymbol{y}, \Delta \boldsymbol{Z})$ leading toward an optimal solution and following a central path defined by the primal and dual constraints and complementarity equation.To make the computation of the direction efficient , the Newtone quations are reduced to the linear system,

$$\boldsymbol{M} \Delta \boldsymbol{y} = \boldsymbol{g},$$

Where the Schur complement matrix $\boldsymbol{M}$ is determined by the constraint matrices $\boldsymbol{A}_i$ and the current point $(\boldsymbol{X}, \boldsymbol{Z})$, and $\boldsymbol{g}$ is defined by current residuals.The IPM repeated ly solves this reduced equation until the it erates atisfies a given convergence tolerance.

It takes $O(mn^3 + m^2n^2)$ operations to compute $\boldsymbol{M}$, which is most expensive part

For each it eration, so we can expect benefit by reducing its computational cost. In many applications of SDP such as the binary code problem, the quadratic assignment problem, and the traveling sales man problem, the matrices $\boldsymbol{A}_i$ and $\boldsymbol{C}$ have identical diagonal block structure. Using the block structure, $\boldsymbol{M}$ can be expanded to

$$\boldsymbol{M} = \sum_{j=1}^{p} \boldsymbol{M}_j,$$

where p is the number of diagonal blocks and matrix $\boldsymbol{M}_j$ is associated with the j-th con- straint block. If some constraint blocks make in significant or detrimental contributions to finding these arch direction, we may be able to ignore the corresponding $\boldsymbol{M}_j$ when we compute $\boldsymbol{M}$. We call such block s *in active*. Similar to the previous problems, it is critical to determine which constraint blocks can be ignored while still guaranteeing that the it eration converges to the optimal solution.

The author explain how constrain treduction can be applied to IPM for SDP problems and propose a basic predictor-correct or algorithm with constraint reduction. We demonstrate its performance by experiments with test problems. The author develops a new predictor-correct or algorithm with adaptive criteria to determine *inactive* constraint blocks. We verify the correctness softhe criteria by proving the global convergence of the proposed algorithm. Its polynomial complex it y is also verified to be $O(n\ln(\varrho_0/\varrho))$, where ϱ_0 is an initial residual and ϱ is a required tolerance.