JORC (2013) 1:199–215
DOI 10.1007/s40305-013-0008-9

Conjugate Decomposition and Its Applications

Li-Ping Wang · Jinyun Yuan

Received: 23 December 2012 / Revised: 25 February 2013 / Accepted: 28 February 2013 /
Published online: 20 March 2013

Abstract The conjugate decomposition (CD), which was given for symmetric and positive definite matrices implicitly based on the conjugate gradient method, is generalized to every $m \times n$ matrix. The conjugate decomposition keeps some SVD properties, but loses uniqueness and part of orthogonal projection property. From the computational point of view, the conjugate decomposition is much cheaper than the SVD. To illustrate the feasibility of the CD, some application examples are given. Finally, the application of the conjugate decomposition in frequency estimate is given with comparison of the SVD and FFT. The numerical results are promising.

Keywords Singular value decomposition · Conjugate decomposition · Generalized inverse · Least squares solution · Projection · Orthogonal projection · Frequency estimate · FFT

1 Introduction

It is well-known that the conjugate gradient method is one of the most popular methods for solving symmetric and positive definite systems of linear equations [3, 6, 8, 15]. All literature looks at the conjugate gradient method as iterative method

The work was partially supported by the CNPq, CAPES, and Fundação Araucária, Brazil, and NSFC11001128, China.

L.-P. Wang
Department of Mathematics, College of Sciences, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, P.R. China
e-mail: wlpmath@yahoo.com.cn

J. Yuan (✉)
Departamento de Matemática – UFPR, Centro Politénico, CP: 19.081, Curitiba, PR, 81531-980, Brazil
e-mail: jin@ufpr.br

from optimization point of view [3, 6, 12, 18]. We just apply the conjugate gradient method to solve large sparse linear systems $Ax = b$ with excellent numerical behavior. Now we can look at the conjugate gradient method purely from the linear algebra point of view. From numerical point of view, the step size α_k in the conjugate gradient method is somehow an approximation of the eigenvalues of A [1] which motivates us to think something similar to singular value decomposition.

Let A be symmetric and positive definite matrix. Suppose that $p_1, p_2, \cdots, p_n$ are conjugate gradient vectors of A which means

$$p_i^T A p_j = \begin{cases} 0, & \text{if } i \neq j \\ > 0 & \text{if } i = j. \end{cases} \tag{1}$$

It is easy to verify that $p_1, p_2, \cdots, p_n$ are linearly independent, further more, form a basis of the vector space R^n. Let us construct a matrix P whose columns are $p_1, p_2, \cdots, p_n$. It is obvious that P is nonsingular. Then, we can find [8, 17] that

$$P^T A P = \operatorname{diag}\big(p_1^T A p_1, p_2^T A p_2, \cdots, p_n^T A p_n\big) \equiv D \geqslant 0, \tag{2}$$

and

$$A^{-1} = P D^{-1} P^T = \sum_{i=1}^{n} \frac{p_i p_i^T}{p_i^T A p_i}, \tag{3}$$

which are similar to the form of the spectral theorem and form of some matrix decomposition.

The matrix decomposition is very powerful and useful for matrix analysis and many real applications [6, 11, 13, 14]. There are several matrix decompositions available in literature [6, 7, 11, 13, 14]. The special case of such matrix decompositions is the singular value decomposition (SVD) [6, 11]. It is well-known that the computational cost of the SVD is very expensive, because it is related to eigenvalues and eigenvectors of $A^T A$ [6]. However, it is relatively cheap to obtain the conjugate gradient vectors for symmetric and positive definite matrices, because we can easily construct n conjugate gradient vectors from every given n linearly independent vector set by three-term recursion given by Hestenes and Stiefel [8] (the comparison of the CPU time for the SVD and the CD can be seen at examples given in Sect. 2). This is the main motivation for us to study the conjugate decomposition (CD) for every given matrix A here.

It is well-known that the Bi-CG method [4] results in the matrix decomposition $\tilde{P}^T A P = D$ for nonsingular A where P and $\tilde{P}$ are nonsingular, and D is diagonal (not necessary nonnegative). While the left conjugate gradient method (LCG) [2, 16, 17] decomposes the matrix A as $P^T A P = T$ where T is triangular, and P is nonsingular. They are all different from the conjugate decomposition because it is similar to the SVD. The similarity of the SVD, the potentially powerful applications (see the last section of this paper) and difference among current matrix decompositions, like Bi-CG and LCG, are the novelties of the conjugate decomposition.

First we reform the conjugate decomposition (CD) for symmetric and positive definite matrix by its conjugate gradient vectors. The most important contribution in this paper is to generalize the CD to rectangular matrices. The CD for rectangular matrix is similar to the singular value decomposition, but loses uniqueness and some orthogonal projection properties. The computational cost of the CD is just $2n$ matrix-vector multiplications which is much less than that of the SVD. Some properties and numerical algorithms of the CD are given. Several simple examples of the CD are given to illustrate its feasibility. Finally we apply the CD to estimate the harmonics in the power system, and give its comparison with the SVD method and the FFT method.

The structure of this short note is as follows. Our main results are given in the next section. We introduce two conjugate decompositions for every $m \times n$ matrix, and study their properties. Some application example of the conjugate decomposition is given as well. In the last section, The CD is applied to power system harmonics estimation.

2 Main Results

In this section, we first reform the conjugate gradient method as conjugate decomposition for symmetric and positive definite matrices, and extend to symmetric and positive semi-definite matrices. Then, we generalize the decomposition to rectangular matrices. Some numerical algorithms and properties are given in this section too.

Lemma 2.1 *Let A be symmetric and positive definite matrix, and $p_1, p_2, \cdots, p_n$ its conjugate gradient vectors. Then, the following decompositions for A and A^{-1} hold:*

$$A = P^{-T} D P^{-1}, \tag{4}$$

and

$$A^{-1} = P D^{-1} P^{T}, \tag{5}$$

where $P = [p_1, p_2, \cdots, p_n]$, $D = \mathrm{diag}(d_1, \cdots, d_n)$ and $d_1 \geqslant d_2 \geqslant \cdots \geqslant d_n > 0$.

The proof is trivial from the conjugate gradient method.

Similarly, we can easily show the following result.

Lemma 2.2 *Let A be a symmetric and positive semi-definite matrix. Then there exists a nonsingular matrix P such that*

$$P^T A P = D, \tag{6}$$

where $D = \mathrm{diag}(d_1, \cdots, d_k, 0, \cdots, 0)$ is diagonal with $d_1 \geqslant d_2 \geqslant \cdots \geqslant d_k > 0$ and $k = \mathrm{rank}(A)$.

Next we shall generalize such decomposition to every matrix.

Theorem 2.3 *Let A be $m \times n$ matrix with rank k. Then, there exist one nonsingular matrix $P \in R^{n\times n}$ and an orthogonal matrix $Q \in R^{m\times m}$ such that*

$$A = Q\Gamma P^{-1}, \tag{7}$$

where $\Gamma \in R^{m\times n}$ is defined as follows

$$\Gamma = \begin{pmatrix} \gamma_1 & 0 & \cdots & 0 & \cdots & 0 \\ r0 & \gamma_2 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \gamma_k & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}_{m\times n}, \tag{8}$$

$\gamma_1 \geqslant \gamma_2 \geqslant \cdots \geqslant \gamma_k > 0$ *and* $\gamma_i = \sqrt{p_i^T A^T A p_i}$ *for* $i = 1, \cdots, k$.

Proof Since $A^T A$ is symmetric and positive semi-definite, by Lemma 2.2, there exists nonsingular matrix $P \in R^{n\times n}$ whose columns are $p_1, \cdots, p_n$ such that

$$A^T A = P^{-T} D P^{-1}, \tag{9}$$

where $D = \mathrm{diag}(p_1^T A^T A p_1, p_2^T A^T A p_2, \cdots, p_k^T A^T A p_k, 0, \cdots, 0)$ with $p_1^T A^T \times A p_1 \geqslant p_2^T A^T A p_2 \geqslant \cdots \geqslant p_k^T A^T A p_k > 0$. Now we define $\gamma_i = \sqrt{p_i^T A^T A p_i} > 0$ for $i = 1, \cdots, k$, and $q_i = \frac{1}{\gamma_i} A p_i$. Then, there are $A p_i = \gamma_i q_i$ for $i = 1, \cdots, k$. Since

$$q_i^T q_j = \frac{1}{\sqrt{p_i^T A^T A p_i}\sqrt{p_j^T A^T A p_j}} p_i^T A^T A p_j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$$

q_i are orthonormal for $i = 1, \cdots, k$. Then, we choose $m - k$ vectors $q_{k+1}, \cdots, q_m$ in R^m such that $q_1, \cdots, q_m$ are orthogonal and $\|q_i\|_2 = 1$ for $i = k+1, \cdots, m$. Then we can rewrite

$$A[p_1, \cdots, p_n] = [q_1, \cdots, q_m]\Gamma, \tag{10}$$

which results in

$$A = Q\Gamma P^{-1}. \tag{11}$$

□

Example 2.4 Assume that A is a 5×3 matrix, randomly generated from a normal distribution with mean 0.6 and standard deviation. Here we use Matlab 7.8 in Thinkpad

PC T410 to do the numerical experiments.

$$A = \begin{pmatrix} 69.2930 & 34.4014 & 103.3326 \\ 35.1039 & -33.1064 & 5.8771 \\ 88.0935 & 105.4856 & 56.7668 \\ 23.7265 & 70.2834 & 52.3648 \\ 26.1993 & 36.1271 & 70.0942 \end{pmatrix}, \tag{12}$$

Matrices Q and P in the CD of A are well defined as Theorem 2.3, and given by

$$Q = \begin{pmatrix} 0.5669 & 0.6084 & 0.0100 & -0.2520 & -0.4949 \\ 0.0247 & 0.4157 & -0.6647 & 0.5748 & 0.2331 \\ 0.6310 & -0.5650 & -0.4699 & -0.2136 & 0.1275 \\ 0.3830 & -0.2544 & 0.4143 & 0.7461 & -0.2455 \\ 0.3650 & 0.2706 & 0.4069 & -0.0611 & 0.7901 \end{pmatrix}, \tag{13}$$

and

$$P = \begin{pmatrix} 1.4583 & 0.0051 & -0.0279 \\ 1.4288 & -0.1647 & 0.0069 \\ 1.9232 & 0.1431 & 0.0166 \end{pmatrix}. \tag{14}$$

Here Γ is 5×3 matrix with the nonzero diagonal elements $(6.1549, 0.1558, 0.0167) \times 10^6$. And $\|A - Q\Gamma P^{-1}\|_2 = 1.1513 \times 10^{-13}$. Denote U, S and V are the singular value decomposition matrices of A, $\|A - USV^T\|_2$ is 1.4673×10^{-13}. The CPU time of computing the CD and the SVD are 1.4952×10^{-4} and 0.0438 seconds respectively.

If randomly given the right hand is $b = [69.8935, 32.6501, 59.0497, 54.7861, 79.8498]^T$, then the least square solutions of $Ax = b$ computed by two methods are the same $x = [0.2053, 0.0019, 0.7425]^T$.

Suppose that $P = [P_1, P_2]$ and $V = P^{-1} = \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix}$ where $P_1, V_1 \in R^{n\times k}$ and $P_2, V_2 \in R^{n\times(n-k)}$. Also suppose that $Q = [Q_1, Q_2]$ where $Q_1 \in R^{m\times k}$ and $Q_2 \in R^{m\times(m-k)}$. Then, the columns of Q_1 and Q_2 form orthonormal basis of $R(A)$ and $N(A^T)$ respectively. Also the columns of V_1 and V_2 form basis, but may not be orthogonal basis of $R(A^T)$ and $N(A)$ respectively. Especially, if we denote $\Gamma_k = \mathrm{diag}\{\gamma_1, \gamma_2, \cdots, \gamma_k\}$, $A = Q_1\Gamma_k V_1^T$ can be considered as the CD economic version which is used more in practical computations.

In the case of $m \leqslant n$, from the computational point of view, the following CD is preferable. The proof is similar to the proof of Theorem 2.3.

Theorem 2.5 *Let A be $m \times n$ matrix with rank k and $m \leqslant n$. Then, there exist one nonsingular matrix $S \in R^{m\times m}$ and an orthogonal matrix $U \in R^{n\times n}$ such that*

$$A = S^{-1}\Gamma U, \tag{15}$$

where $\Gamma \in R^{m\times n}$ is defined by

$$\Gamma = \begin{pmatrix} \gamma_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \gamma_k & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}_{m \times n}, \tag{16}$$

$\gamma_1 \geqslant \gamma_2 \geqslant \cdots \geqslant \gamma_k > 0$, *and* $\gamma_i = \sqrt{s_i^T A A^T s_i}$ *for* $i = 1, \cdots, k$.

Corollary 2.6 *For every* $m \times n$ *matrix* A *with rank* k, *we decompose* A *as*

$$A = \sum_{i=1}^{k} \sqrt{p_i^T A^T A p_i} q_i v_i^T, \tag{17}$$

where p_i *are conjugate vectors of* $A^T A$, $q_i = \frac{1}{\sqrt{p_i^T A^T A p_i}} A p_i$ *and* v_i *are columns of the inverse of* $P = [p_1, \cdots, p_n]$. *Note that the CD is the SVD if the matrix* P *is orthogonal.*

Definition 2.7 For given matrix $A \in R^{m \times n}$, B is called the inverse-$\{1, 2, 3\}$ of A if a matrix $B \in R^{n \times m}$ satisfies the following three equations

$$\begin{cases} ABA = A, \\ BAB = B, \\ (AB)^T = AB. \end{cases} \tag{18}$$

Then, we denote B as $A^{\{1,2,3\}}$.

Definition 2.8 For given matrix $A \in R^{m \times n}$, B is called the inverse-$\{1, 2, 4\}$ of A if a matrix $B \in R^{n \times m}$ satisfies the following three equations

$$\begin{cases} ABA = A, \\ BAB = B, \\ (BA)^T = BA. \end{cases} \tag{19}$$

Then, denote B as $A^{\{1,2,4\}}$.

Theorem 2.9 *For every* $m \times n$ *matrix* A *with rank* k *and* $m \geqslant n$, $A^{\{1,2,3\}}$ *is given by*

$$A^{\{1,2,3\}} = P \Gamma^{\dagger} Q^T, \tag{20}$$

and

$$A^{\{1,2,3\}} = \sum_{i=1}^{k} \frac{1}{\sqrt{p_i^T A^T A p_i}} p_i q_i^T, \tag{21}$$

where $\Gamma^{\dagger}$ *is the generalized inverse of* Γ *given in* (16).

Proof Let $B = P\Gamma^\dagger Q^T$. Then,

$$ABA = Q\Gamma P^{-1} P\Gamma^\dagger Q^T Q\Gamma P^{-1} = Q\Gamma P^{-1} = A,$$
$$BAB = P\Gamma^\dagger Q^T Q\Gamma P^{-1} P\Gamma^\dagger Q^T = P\Gamma^\dagger Q^T = B,$$
$$AB = Q\Gamma P^{-1} P\Gamma^\dagger Q^T = Q\Gamma\Gamma^\dagger Q^T$$

is symmetric. But $BA = P\Gamma^\dagger Q^T A\Gamma P^{-1}$ is not symmetric. By the definition of $A^{\{1,2,3\}}$, the proof is complete. □

Similarly, we can easily prove the following result.

Theorem 2.10 *For every $m \times n$ matrix A with rank k and $m \leqslant n$, $A^{\{1,2,4\}}$ is given by*

$$A^{\{1,2,4\}} = U\Gamma^\dagger S^T, \tag{22}$$

and

$$A^{\{1,2,4\}} = \sum_{i=1}^{k} \frac{1}{\sqrt{s_i^T A^T A s_i}} u_i s_i^T. \tag{23}$$

Corollary 2.11 *$A^{\{1,2,3\}}$ and $A^{\{1,2,4\}}$ have the following projection properties.*

1. *$AA^{\{1,2,3\}}$ is an orthogonal projection onto $R(A)$.*
2. *$A^{\{1,2,3\}}A$ is just projection.*
3. *$A^{\{1,2,4\}}A$ is an orthogonal projection onto $R(A^T)$.*
4. *$AA^{\{1,2,4\}}$ is just projection.*

Proof Since $(AA^{\{1,2,3\}})(AA^{\{1,2,3\}}) = AA^{\{1,2,3\}}$ and $AA^{\{1,2,3\}}$ is symmetric, $AA^{\{1,2,3\}}$ is orthogonal projection. Note that $(A^{\{1,2,3\}}A)(A^{\{1,2,3\}}A) = A^{\{1,2,3\}}A$, and $A^{\{1,2,3\}}A$ is nonsymmetric, and hence, $A^{\{1,2,3\}}A$ is a projection. Now $A^{\{1,2,4\}}A$ is orthogonal projection which follows from $A^{\{1,2,4\}}AA^{\{1,2,4\}}A = A^{\{1,2,4\}}A$ and $(A^{\{1,2,4\}}A)^T = A^{\{1,2,4\}}A$. Since $AA^{\{1,2,4\}}$ is not symmetric, and $AA^{\{1,2,4\}}AA^{\{1,2,4\}} = AA^{\{1,2,4\}}$, $AA^{\{1,2,4\}}$ is only projection. □

Next we shall give some feasible and cheap numerical algorithm to compute $A^{\{1,2,4\}}$ and $A^{\{1,2,3\}}$ respectively. Without losing the generality, we shall normalize vectors p_i for $i = 1, \cdots, k$ which requires $\|p_i\|_2 = 1$ for $i = 1, \cdots, n$.

Algorithm 2.1

1. *Given any $m \times n$ matrix A with rank$(A) = k$, choose one $r_1 \neq 0$. Set $p_1 = r_1/\|r_1\|_2$*
2. *For $i = 2, \cdots, k$*
   2.1. *Compute p_i of $A^T A$ by the conjugate gradient method*

$$q = Ap_{i-1},$$
$$r_i = r_{i-1} - \alpha A^T q,$$

$$p_i = r_i + \beta p_{i-1},$$

$$p_i = p_i/\|p_i\|_2.$$

2.2. *calculate* $q_i = \frac{1}{\sqrt{q^T q}} q$

3. *Choose* p_j *orthonormal to all* p_i *for* $i = 1, \cdots, k$ *and* $j = k+1, \cdots, n$ *in null space of* A.
4. *Choose* q_j *orthonormal to* q_i *with* $\|q_j\|_2 = 1$ *in* R^m.

Algorithm 2.2

1. *Given any* $m \times n$ *matrix* A *with rank*$(A) = k$, *choose one* $r_1 \neq 0$. *Set* $p_1 = r_1/\|r_1\|_2$.
2. *For* $i = 2, \cdots, k$

   2.1. *Compute* p_i *of* AA^T *by the conjugate gradient method*

$$q = A^T p_{i-1},$$

$$r_i = r_{i-1} - \alpha A q,$$

$$p_i = r_i + \beta p_{i-1},$$

$$p_i = p_i/\|p_i\|_2.$$

   2.2. *calculate* $q_i = \frac{1}{\sqrt{q^T q}} q$

3. *Choose* p_j *orthonormal to all* p_i *for* $i = 1, \cdots, k$ *and* $j = k+1, \cdots, m$ *in null space of* A.
4. *Choose* q_j *orthonormal to* q_i *with* $\|q_j\|_2 = 1$ *in* R^n.

Remark 1 Algorithm to calculate the CDs or $A^{\{1,2,3\}}$ (or $A^{\{1,2,3\}}$) is much easier and cheaper than the algorithm to compute the SVD. In this sense, the CD probably is more feasible and practical than the SVD for some applications.

We further note that there might be stability problem for the CDs with current computational technology.

Example 2.12 Let u be column vector in R^n. Then,

$$u = [q_1, q_2, \cdots, q_n] \begin{pmatrix} \sqrt{u^T u} \\ 0 \\ \vdots \\ 0 \end{pmatrix} 1,$$

where

$$q_1 = \frac{u}{\sqrt{u^T u}},$$

Table 1 Problem

No.	Group and name	Problem kind	Size(double)	Date
1	NYPA/Maragal_5	Least square problem	4654×3320	2008
2	Meszaros/gas11	Linear programming problem	459×862	2004

Table 2 Numerical results of CD and SVD for problem NYPA/Maragal_5

		CPU time (s)	$\|AA_k^\dagger A - A\|_F$	$\|A_k^\dagger AA_k^\dagger - A\|_F$	$\|(AA_k^\dagger)^T - AA_k^\dagger\|_F$
$k=5$	CD	0.1267	112.0763	1.9477e−015	2.1922e−015
	SVD	0.7982	111.4893	7.2112e−014	3.1549e−012
$k=10$	CD	0.0388	110.7114	1.3178e−012	3.0886e−015
	SVD	0.2132	109.4627	1.6732e−012	7.1045e−011

Note: $\|\cdot\|_F$ denote the Frobenius norm, evaluating the decomposition error

Table 3 Numerical results of CD and SVD for Problem Meszaros/gas11

		CPU time (s)	$\|AA_k^\dagger A - A\|_F$	$\|A_k^\dagger AA_k^\dagger - A\|_F$	$\|(A_k^\dagger A)^T - A_k^\dagger A\|_F$
$k=5$	CD	4.7855e−004	604.3581	2.4479e−017	2.8246e−016
	SVD	0.0582	465.9531	2.7152e−016	5.3384e−013
$k=10$	CD	9.9884e−004	587.6938	8.9382e−017	9.0305e−016
	SVD	0.0350	419.8070	4.3707e−014	7.9358e−011

Note: $\|\cdot\|_F$ denote the Frobenius norm, evaluating the decomposition error

q_i are orthogonal to q_1 for $i = 2, \cdots, n$. Therefore,

$$u^\dagger = u^W = 1\left[\frac{1}{\sqrt{u^T u}}, 0, \cdots, 0\right]\begin{pmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_n^T \end{pmatrix}.$$

Note that in this case, the CD is the same as the SVD for u.

Remark 2 In practical applications, such as linear least square problems, complete matrix decomposition is unnecessary. For different k, we present numerical comparisons of the CD and the SVD for the problems whose properties are listed in Table 1 from http://cise.ufl.edu/research/sparse/matrices/.

Tables 2 and 3 show the promising efficiency of the incomplete CD.

Remark 3 Analogous to "Singular Value" of the SVD, we define the $\gamma_1, \gamma_2, \cdots, \gamma_k$ "Conjugate Values" of the CD. Interestingly, Conjugate Values roughly simulate the

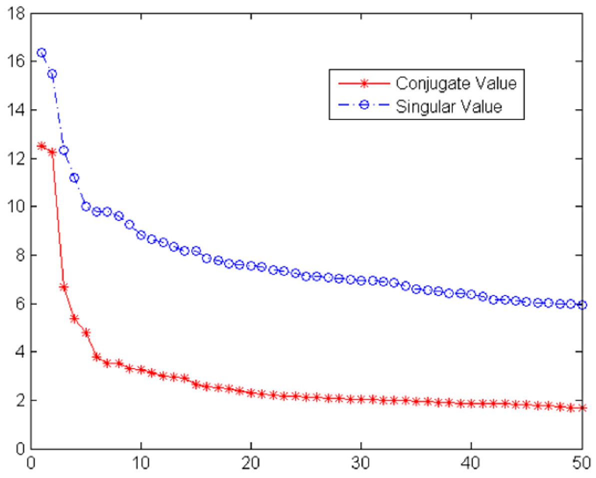


Fig. 1 Conjugate and singular values distribution

variability of Singular Values which is our future research issue. For example, we give the largest 50 Conjugate Values and Singular Values of problem NYPA/Maragal_5 in Fig. 1. This property might be useful in some applications such as pattern recognition, signal processing, etc.

Consider least squares problem

$$\min \|Ax - b\|_2, \tag{24}$$

where $A \in R^{m\times n}$ with full rank.

Theorem 2.13 *Suppose that $A = Q\Gamma P^{-1} \in R^{m\times n}$ with full rank and $m \geqslant n$. Then, the solution of problem* (24) *is given by*

$$x = A^{\{1,2,3\}}b = \sum_{i=1}^{n} \frac{q_i^T b}{\sqrt{p_i^T A^T A p_i}} p_i. \tag{25}$$

If $m \leqslant n$, then the minimal norm solution is given by

$$x = A^{\{1,2,4\}}b = \sum_{i=1}^{k} \frac{p_i^T b}{\sqrt{p_i^T A A^T p_i}} q_i. \tag{26}$$

Proof First we suppose that $m \geqslant n$. By Lemma 2.1, the solution of the problem (24) is given by $x = (A^T A)^{-1} A^T b$. By the CD, we can obtain that

$$x = \left(P^{-T}\Gamma^T Q^T Q\Gamma P^{-1}\right)^{-1} P^{-T}\Gamma^T Q^T b = P\left(\Gamma^T\Gamma\right)^{-1}\Gamma^T Q^T b$$

which results in

$$x = P\Gamma^{\dagger} Q^T b = A^{\{1,2,3\}}b.$$

Table 4 The SVD and CD to solve least square problem NYPA/Maragal_5

		CPU time (s)	$\|b - AA_k^{\dagger}b\|_F$
$k=5$	CD	0.0173	5.6921
	SVD	0.1474	9.1399
$k=10$	CD	0.0400	3.8698
	SVD	0.2258	8.4590
$k=20$	CD	0.0782	1.7276
	SVD	0.5210	8.0161
$k=1000$	CD	11.857	7.35×10^{-4}
	SVD	1.01×10^{3}	7.0311

Hence,

$$x=\sum_{i=1}^{n}\frac{p_i^T A^T b}{p_i^T A^T A p_i}p_i=\sum_{i=1}^{n}\frac{q_i^T b}{\sqrt{p_i^T A^T A p_i}}p_i. \tag{27}$$

Now we assume that $m \leqslant n$. In this case, the minimal norm solution is given by $x = A^T(AA^T)^{-1}b$. By the CD, we have

$$\begin{aligned}x &= A^T\left(AA^T\right)^{-1}b = Q\Gamma^T P^{-1}\left(P^{-T}\Gamma Q^T Q\Gamma^T P^{-1}\right)^{-1}b\\ &= Q\Gamma^T\left(\Gamma\Gamma^T\right)^{-1}P^T b = Q\Gamma^{\dagger}P^T b = A^{\{1,2,4\}}b = \sum_{i=1}^{k}\frac{p_i^T b}{\sqrt{p_i^T AA^T p_i}}q_i,\end{aligned}$$

and this completes the rest of the proof. □

Example 1 In general, for any arbitrary $m \times n$ matrix, the solution of the least squares problem (24) is given by

$$x=\begin{cases}A^{\{1,2,3\}}b & \text{if } m \geqslant n,\\ A^{\{1,2,4\}}b & \text{if } m \leqslant n.\end{cases} \tag{28}$$

Table 4 is the results to solve overdetermined problem NYPA/Maragal_ using the incomplete CD and the incomplete SVD. With the same k, the CD gets better evaluations of least square solution than the SVD.

3 Harmonics Estimation

The quality of the delivered energy is deteriorated, the energy loses are increased as well as the reliability of a power system is decreased because of wide spectrum of harmonic components generated by modern frequency power converters [5]. Then, the parameter estimation of the components is very important for control and protection tasks. Even the FFT is computational efficient with reasonable results for spectrum estimation of discretely sampled processes, it has several performance limitations, one of the most prominent limitation is that of frequency resolution, that is, the ability to distinguish the spectral responses of two or more signals. To avoid such limitations, the SVD was used to do harmonics estimation. The numerical tests for the SVD method displayed the superiority of the approach for signals buried in noise to the FFT method as a very versatile and efficient tool for detection and location of all higher harmonics existing in the system. However, the computational cost of the SVD is much more complex and expensive than the FFT [10]. For the limitation of the SVD, we hire CD to do the same work.

In power system harmonics estimation, an overdetermined linear system is established in autoregressive (AR) model. Consider the signal waveform [9]

$$\begin{aligned} x(t) &= 100\cos 2\pi 40t + 50\cos 2\pi 217t + 40\cos 2\pi 760t \\ &\quad + 40\cos 2\pi 1000t + \kappa_s e(t), \end{aligned} \tag{29}$$

where $e(t)$ is a white noise of zero mean and variance equal to 1. The signal waveform contains the basic component 40 Hz, two higher harmonics 1000 Hz, 760 Hz, and one interharmonics 217 Hz. The sample period is 0.4 ms, the number of samples $n = 150$. Here $\kappa_s = 40$ and the order $l = 70$ in impulse response $H(z) = 1 - \sum_{i=1}^{l} h_i z^{-i}$ of AR model. Estimation of harmonics is equivalent to solving overdetermined system [9, 10]

$$Ah = b, \tag{30}$$

where

$$A = \begin{pmatrix} x_l & x_{l-1} & \cdots & x_1 \\ x_{l+1} & x_l & \cdots & x_2 \\ \cdots & \cdots & \cdots & \cdots \\ x_{n-1} & x_{n-2} & \cdots & x_{n-l} \\ x_2 & x_3 & \cdots & x_{l+1} \\ x_3 & x_4 & \cdots & x_{l+2} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n-l+1} & x_{n-l+2} & \cdots & x_n \end{pmatrix}, \quad h = \begin{pmatrix} h_1 \\ h_2 \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ h_l \end{pmatrix}, \quad \text{and} \quad b = \begin{pmatrix} x_{l+1} \\ x_{l+2} \\ \cdots \\ x_n \\ x_1 \\ x_2 \\ \cdots \\ x_{n-l} \end{pmatrix}. \tag{31}$$

Both the CD and SVD can be used to solve (30). Actually, the fast Fourier transform (FFT) is conventional approach in spectral estimation. Then three measured waveforms with the same sample number and period are comparatively presented. Similar to only dominant singular values contributing to harmonics signals [9], conjugate value variabilities are useful in measuring waveform. At noiseless measurement, the distributions of conjugate values and singular values for A are presented in Fig. 2. Sharp gap happens between the 9th and 10th values. The first 9 highest values correspond to fundamental frequency and harmonics, while the remaining values are close to each other which correspond to noise. The relation between two kinds of values are largely dependent on the noise to signal ratio. Theoretically, we can remove them or reduce to economic decompositions. Figure 3(a) is the CD waveform with the 10

Fig. 2 Distribution of conjugate and singular values for A

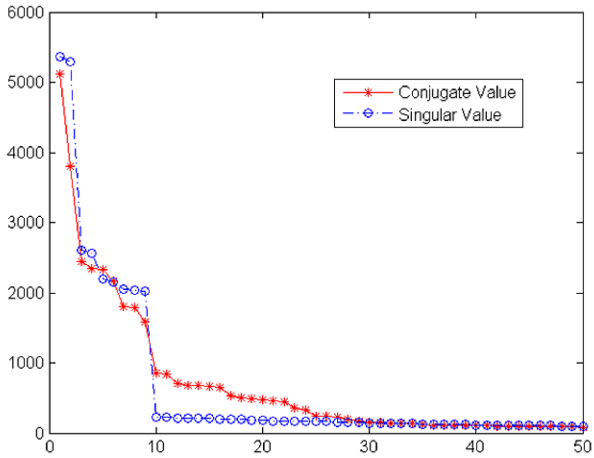
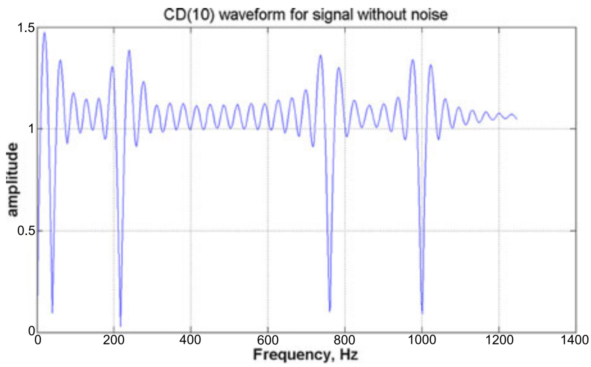
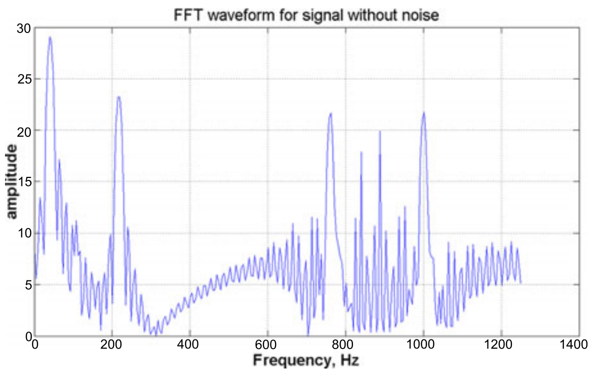


Fig. 3 $H(z)$ measured model



(a) CD(10) waveform



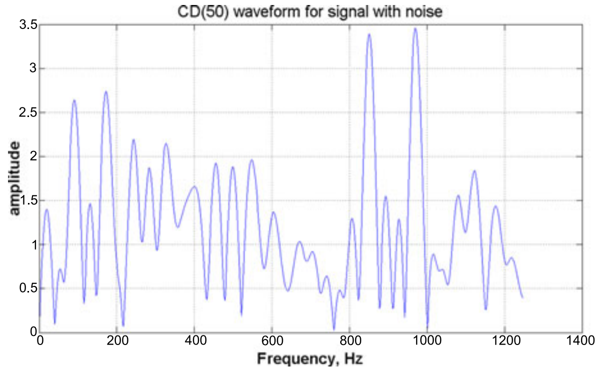
(b) FFT waveform

largest conjugate values and Fig. 3(b) is the FFT. The zeros or the points closest zero in CD waveform determine the exact frequencies, as well as the peak values in the FFT waveform. The CD and FFT clearly recognize all the harmonics (40 Hz, 760 Hz

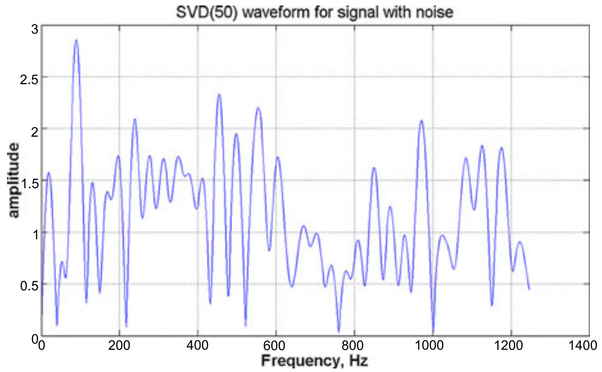
Table 5 The measured results by the CD(50), SVD(50) and FFT

	40	217	760	1000	Time (s)
CD	39.5	217	759.5	1000.5	1.05×10^{-3}
SVD	39.5	217	759.5	1000.5	4.63×10^{-2}
FFT	39	219	762	1001	7.62×10^{-4}

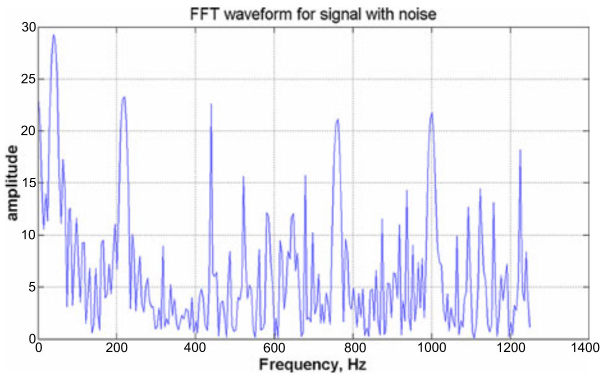
Fig. 4 $H(z)$ measured model



(a) CD(50) waveform

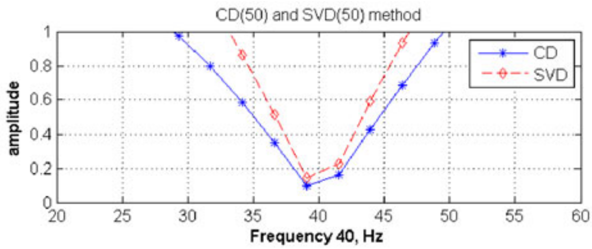


(a) SVD(50) waveform

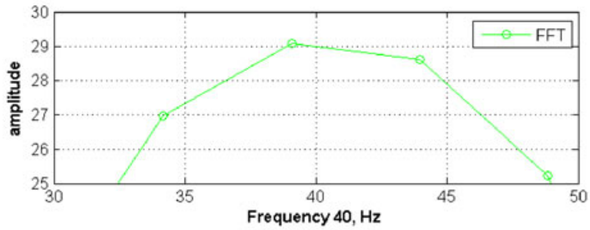


(b) FFT waveform

Fig. 5 Magnitude frequency 40

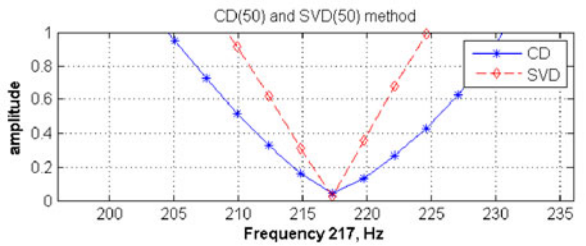


(a) The CD(50) and SVD(50)

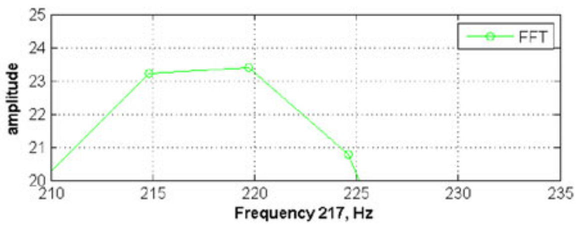


(b) FFT

Fig. 6 Magnitude frequency 217



(a) CD(50) and SVD(50)

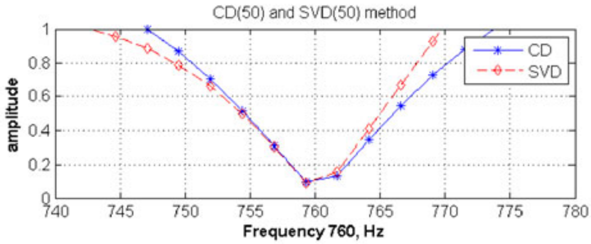


(b) FFT

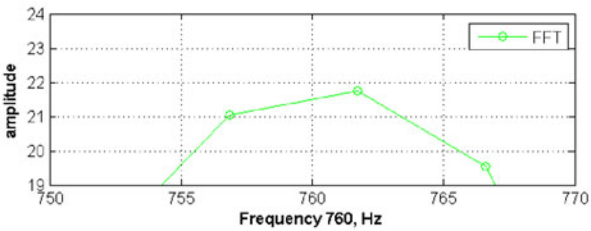
and 1000 Hz) and interharmonics (217 Hz). Unfortunately, the 10 largest singular values can not well construct solution to (30), then the SVD(10) fails.

In practically noisy environment, the measured waveforms of CD(50), the SVD(50) and FFT are presented. The CD(50) and SVD(50) estimate all harmonics and inter harmonics more accurate than the FFT (Table 5). Noise disturbing causes

Fig. 7 Magnitude frequency 760

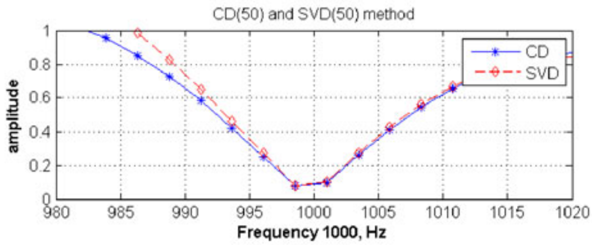


(a) CD(50) and SVD(50)

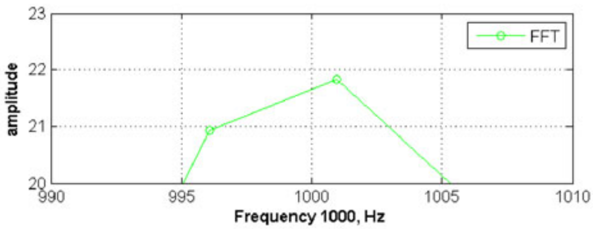


(b) FFT

Fig. 8 Magnitude frequency 1000



(a) CD(50) and SVD(50)



(b) FFT

the FFT spectrum leakage around 560 Hz (Fig. 4(b)). Comfortingly, the CD(50) is faster than the SVD(50) in spite of that slower than the FFT (Table 5). More details can be seen in Magnitude Figs. 5, 6, 7 and 8.

Acknowledgements Authors like to give their thanks to the referees for their helpful comments and suggestions to improve the quality of the paper. Authors also thank Professors Yu-Hong Dai, José Mario Martinez, and Yaxiang Yuan for their helpful discussions and suggestions which help us to improve the quality of this article and further research issues.

References

[1] Dai, Y.-H.: Private communication (2006)
[2] Dai, Y.-H., Yuan, J.Y.: Study on semi-conjugate direction methods for non-symmetric systems. Int. J. Numer. Methods Eng. **60**, 1383–1399 (2004)
[3] Demmel, J.W.: Applied Numerical Linear Algebra. SIAM, Philadelphia (1997)
[4] Fletcher, R.: Conjugate gradient method for indefinite systems. Numer. Anal. Lect. Notes Math. **506**, 73–89 (1976)
[5] Gentile, G., Rotondale, N., Tursini, M., Franceschini, G., Tassoni, C.: Fault operation of inverter-fed induction motor. In: Proceedings of the 31st Universities Power Engineering Conference, Iraklion, Greece, vol. 2, pp. 357–360 (1996)
[6] Golub, G.H., Van Loan, C.: Matrix Computation, 3rd edn. John Hopkins University Press, Baltimore (1996)
[7] Golub, G.H., Yuan, J.Y.: ST: symmetric-triangular decomposition and its applications. Part I: Theorems and algorithms. BIT **42**, 814–822 (2002)
[8] Hestens, M.R., Stiefel, E.: Methods of conjugate gradients for solving linear systems. J. Res. Natl. Bur. Stand. **49**(6), 409–436 (1952)
[9] Osowski, S.: SVD technique for estimation of harmonic components in a power system: a statistical approach. IEE Proc., Gener. Transm. Distrib. **141**(5), 473–479 (1994)
[10] Lobos, T., Kozina, T., Koglin, H.J.: Power system harmonics estimation using linear least squares method and SVD. IEE Proc., Gener. Transm. Distrib. **148**(6), 567–572 (2001)
[11] Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press, Cambridge (1985)
[12] Nocedal, J., Wright, S.J.: Numerical Optimization. Springer, Berlin (1999)
[13] Stewart, G.W.: Matrix Algorithms, Volume I: Basic Decompositions. SIAM, Philadelphia (1998)
[14] Stewart, G.W.: Matrix Algorithms, Volume II: Eigensystems. SIAM, Philadelphia (2001)
[15] Trefethen, L.N., Bau, D. III: Numerical Linear Algebra. SIAM, Philadelphia (1997)
[16] Wang, L.-P., Dai, Y.-H.: Left conjugate gradient method for non-Hermitian linear systems. Numer. Linear Algebra Appl. **15**, 891–909 (2008)
[17] Yuan, J.Y., Golub, G.H., Plemmons, R.J., CECILIO, W.A.G.: Semi-conjugate direction methods for real positive definite systems. BIT Numer. Math. **44**(1), 189–207 (2004)
[18] Yuan, Y.: Computational Methods for Nonlinear Optimization. Science Press, Beijing (2008)