A Theoretical Study on the Orientation Problem in Linear Wireless Sensor Networks*

Jianghong Han[1,2], Xu Ding[1,2], Lei Shi[1,2], Dong Han[3], and Zhenchun Wei[1,2]

[1] School of Computer Science and Information,
Hefei University of Technology, Hefei 230009, China
[2] Engineering Research Center of Safety Critical Industrial Measurement
and Control Technology, Ministry of Education, Hefei 230009, China
[3] Department of Computer Science, University of Houston, Houston, TX 77004, USA

Abstract. A theoretical approach of acquiring arrival angles of signals sensed by sensor nodes in linear wireless sensor networks is introduced. The arrival angles of signals can be obtained by the estimation of signal covariance matrices. In this article, firstly, the existence of the solution to the estimation problem is studied intensively. Later on, the solution to this problem of estimating real-valued covariance matrices is discussed by the approach of maximum-likelihood estimation. Finally, this approach is expanded to the realm of complex-valued covariance matrices.

Keywords: arrival angles of signals, estimation of covariance matrices, maximum-likelihood estimation.

1 Introduction

Wireless sensor networks (WSNs) consisted of huge amount of sensor nodes and base station(s) are capable of performing numerous unmanned tasks in extreme environments, such as volcano areas, toxic areas, underwater and underground etc. WSNs play a very important role in many aspects of our modern society. For instance, with the aid of wireless sensor network, we can make our weather forecast more precise. In situations involving emergency services, such as poisonous gas leaks, wireless sensor networks will reduce our costs on locating and rescuing persons in danger.

Linear wireless sensor networks (LWSNs) are a special family of wireless sensor networks with regard to linear network topology. Compared with normal wireless sensor networks, LWSNs exhibit lower complexity. Even though simple in topology, LWSNs possess numerous practical applications, such as monitoring public transportations, oil pipes, factories and plants. Nowadays, intensive researches in the field of WSNs focus on Network protocols, such as routing protocols, MAC protocols and cross-layer protocols, etc., which aim to alleviate the energy overhead of sensor nodes

* This paper is sponsored by National Natural Science Foundation of China(NSFC) with Grant No. 60873003, No.60873195 and Ph. D. Programs Foundation of Ministry of Education of China with Grant No. 2010JYBS0762.

X. Wang et al. (Eds.): WASA 2012, LNCS 7405, pp. 722–732, 2012.

and meet the real-time requirements in data transmission[1-6]. However, this article will primarily introduce a theoretical study of the estimation of signal covariance matrix in linear sensor networks which will be helpful in dealing with determining the direction of events popping up in the inspecting field of a LWSN.

The remainder of this article is organized as follows: in section 2, the covariance matrix estimation problem will be briefly introduced as well as the mathematical modeling of the signal covariance matrix of LWSNs. What is more, the underlying relationship between the direction determination problem and the estimation of signal covariance matrix is also revealed given this model. Section 3 will look into the existence of solutions to the estimation problems referring to the maximum-likelihood estimation. In section 4, the attention will be paid to how to obtain such a solution to the estimation problem. Additionally, the result of the estimation problem with respect to real-valued covariance matrices will be broaden into the realm of complex-valued covariance matrices which have a close relationship to orientation problems in LWSNs. Section 5 will conclude this article.

2 Brief Introduction to the Estimation of Covariance Matrix

In the area of statistics, the estimation of a covariance matrix[7-10] is to approximately determine the unknown covariance matrix C of an M-dimension multivariate random variable R given a series of $\mathrm{x}_1, \mathrm{x}_2, \dots, \mathrm{x}_\mathrm{N}$. Each x_i is an M-dimension vector drawn from the multivariate distribution of which the probability density function is $\mathrm{p}(\mathrm{x}_\mathrm{i})$. And the covariance matrix C in calculated by $\mathrm{E}[(\mathrm{R}-\mathrm{E}(\mathrm{R}))(\mathrm{R}-\mathrm{E}(\mathrm{R}))^\mathrm{H}]$, where $\mathrm{E}(\cdot)$is an expectation, and $(\cdot)^\mathrm{H}$ denotes the conjugate transpose of a matrix.

Assume that N events of interest burst out in the deploying area of the LWSN, and are sensed by M sensor node. Recall that these sensors are arranged in a linear fashion as shown in fig.1.

Suppose that the signals are narrow-banded with certain known frequency f in advance, and the sensors are equally spaced with respect to each other. It is also assumed that these signals propagate over the distance long enough to make sure that the N received signals by all M sensors are parallel to each other as shown in fig.2.

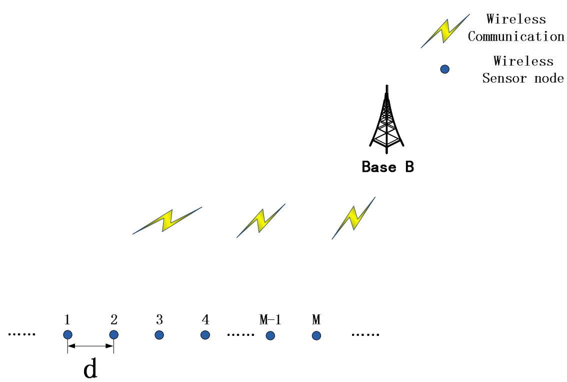

Fig. 1. Wireless Sensor Networks deployed in a linear fashion

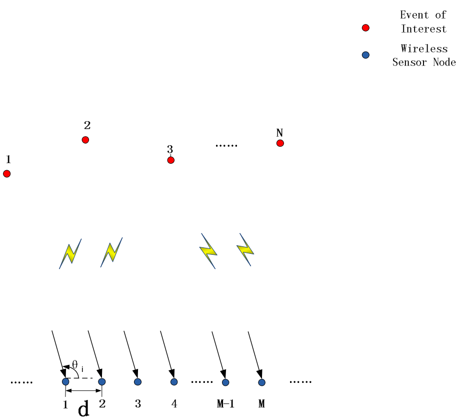

Fig. 2. N events of interest sensed by M sensor nodes

The N signals can be expressed by a M-by-N array manifold matrix G of the form

$$G = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \exp\left(-\frac{j2\pi d}{\lambda}\sin\theta_1\right) & \exp\left(-\frac{j2\pi d}{\lambda}\sin\theta_2\right) & \cdots & \exp\left(-\frac{j2\pi d}{\lambda}\sin\theta_N\right) \\ \vdots & \vdots & \vdots & \vdots \\ \exp\left(-\frac{j(M-1)2\pi d}{\lambda}\sin\theta_1\right) & \exp\left(-\frac{j(M-1)2\pi d}{\lambda}\sin\theta_2\right) & \cdots & \exp\left(-\frac{j(M-1)2\pi d}{\lambda}\sin\theta_N\right) \end{bmatrix}_{M\times N}$$

where λ denotes the wavelength of the signal, d denotes the distance between two sensor nodes, and θ_i denotes the arrival angle of the i^{th} event with respect to the line of sensor nodes. If the received signals are interfered by some additive noise with zero mean and σ_n^2 variance Gaussian distribution which is uncorrelated with signals, the covariance matrix C of the received signals is of the form:

$$C = C_N + GPG^H \tag{1}$$

where C_N denotes the covariance matrix of M-dimension additive noise with zero mean and σ_n^2 variance Gaussian distribution, P denotes the correlation matrix between signals. Furthermore, it the N signals are also independent to each other, (1) can be cast as

$$C = \sigma_n^2 I_M + [g_1 \quad g_2 \quad \cdots \quad g_N] \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_N^2 \end{bmatrix} [g_1 \quad g_2 \quad \cdots \quad g_N]^H = \sigma_n^2 I_M + \sum_{i=1}^{N} \sigma_i^2 g_i g_i^H \tag{2}$$

where $g_i = \left[1 \quad \exp\left(-\frac{j2\pi d}{\lambda}\sin\theta_1\right) \quad \cdots \quad \exp\left(-\frac{j(M-1)2\pi d}{\lambda}\sin\theta_N\right)\right]^T$, σ_n^2 is the noise power, σ_i^2 is the i^{th} signal power.

Till now, it is clear that the covariance matrix C can be parameterized by σ_n^2, $\{\sigma_i^2\}$, and $\{\theta_i\}$, i.e.,

$$C = C(\sigma_n^2, \{\sigma_i^2\}, \{\theta_i\}) \tag{3}$$

From (3), it is obvious that the covariance matrix conveys knowledge about the arrival angle θ_i of the i^{th} signal in that C is a function of $\{\theta_i\}$ which will help us determine the direction of this signal. In the following sections, our discussion is mainly focused on how to estimating covariance matrices with certain special structure which is also possessed by C.

3 Estimation of Real-Valued Covariance Matrices with Certain Structure

The covariance matrix C is a complex valued toeplitz (please refer to the appendix for detailed information) matrix as is shown in (2). What is more, C is also a hermitian matrix since C is equal to C^H. However, even if C is such a structured matrix, it is still not easy to estimate C in that C is complex valued. Therefore, we first introduce the method which can be used to estimate some structured real valued matrices, and then the result will be extended to the realm of complex valued matrices.

3.1 Estimating the Covariance Matrices through Maximum-Likelihood Method

Suppose that the N real valued samples $x_1, x_2, \ldots, x_N$ are drawn from an M-dimension Gaussian distribution with zero mean and covariance matrix C. Therefore, the probability density function for each x_i is

$$p(x_i) = (2\pi)^{-\frac{M}{2}} \det(C)^{-\frac{1}{2}} exp\left(-\frac{x_i{}^T C^{-1} x_i}{2}\right) \tag{4}$$

If all these samples are independent to each other, the joint probability density function of these N samples is

$$p(x_1, x_2, \ldots, x_N) = (2\pi)^{-\frac{MN}{2}} \det(C)^{-\frac{N}{2}} \exp\left(-\sum_{i=1}^{N} \frac{x_i{}^T C^{-1} x_i}{2}\right) \tag{5}$$

In (4) and (5), the covariance matrix C is unknown and to be estimated, however, we presume that C is with certain structure.

The log-likelihood function of C is

$$\begin{aligned} \ln\big(L(C;\{x_i\})\big) &= \ln\big(p(x_1, x_2, \ldots, x_N)\big) = -\frac{MN}{2}\ln 2\pi - \frac{N}{2}\ln\det(C) - \frac{1}{2}\sum_{i=1}^{N} x_i{}^T C^{-1} x_i \\ &= \frac{MN}{2}\ln 2\pi - \frac{N}{2}\ln\det(C) - \frac{1}{2}\mathrm{tr}\left(C^{-1}\sum_{i=1}^{N} x_i x_i{}^T\right) \end{aligned} \tag{6}$$

The proof of the second equation can be found in Appendix.

Instead of maximizing the log-likelihood function stated in (6), it is equivalent to maximize the function

$$L'(C;\{x_i\}) = -\ln\det(C) - \mathrm{tr}\left(C^{-1}\frac{1}{N}\sum_{i=1}^{N} x_i x_i^{T}\right) \tag{7}$$

Our estimator of C is $\hat{C} = \mathrm{argmax}(L'(C;\{x_i\}))$.

3.2 The Existence of the Solution to the Maximum-Likelihood Problem of Certain Structured Covariance Matrices

In previous section, we have pointed out the estimation problem we are about to solve. In this section, a further study on the existence of solutions to this estimation problem will be carried out before discussing how to obtain such solutions. The real valued nonnegative-definite symmetric matrices will be taken into consideration because the observed samples will form such structured matrices.

Let $S = \frac{1}{N}\sum_{i=1}^{N} x_i x_i^{T}$, and substitute S into (7), we have

$$L'(C;S) = -\ln\det(C) - \mathrm{tr}(C^{-1}S) \tag{8}$$

Obviously, S is a nonnegative definite symmetric matrix, and our estimation problem can be formulated as a optimization problem

Maximize: $L'(C;S) = -\ln\det(C) - \mathrm{tr}(C^{-1}S)$

Subject to: C belongs to the set of nonnegative definite matrices

Before discussing the optimization problem, it is presumed that S is a positive definite matrix for every sample x_i we acquire only possesses zero valued entries if S is nonnegative definite but not positive definite, which will make the problem extremely difficult to deal with.

To prove the existence of solutions, we first look into the value of the objective function. We can prove that the objective function has an upper bound, and when $\det(C)$ goes to zero, the value of the function tends to be minus infinity. To prove this, we need a lemma first.

Lemma 1: if A and B are two positive definite symmetric matrices, there exists one unitary matrix U which will shoe these two matrices into diagonal matrices simultaneously through congruent transformation. (The proof of this lemma can be found in Appendix)

Theorem 1: the value of the objective function in above optimization problem has an upper bound, and tends to be minus infinity while $\det(C)$ goes to zero.

Proof: Firstly, suppose that C is positive definite, which means that $\det(C)$ is greater than zero. Because C and S are both real valued positive definite symmetric matrices, we know from Schur Theorem that there exists unitary matrices U_1 and U_2 which will make C and S congruent to two diagonal matrices. Moreover, from lemma 1, there will be one unitary matrix U which could shape C and S congruent to diagonal matrices simultaneously, i.e.,

$$\mathrm{U^TCU} = \Lambda_C = \begin{bmatrix} c_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_M \end{bmatrix}, \text{ and } \mathrm{U^TSU} = \Lambda_S = \begin{bmatrix} s_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & s_M \end{bmatrix}$$

Therefore, $\mathrm{tr}(C^{-1}S)$ can be rewritten as

$$\mathrm{tr}(C^{-1}S) = \mathrm{tr}(U(U^TCU)^{-1}U^TSUU^T) = \mathrm{tr}(U\Lambda_C^{-1}\Lambda_S U^T) = \mathrm{tr}(\Lambda_S\Lambda_C^{-1}) = \sum_{i=1}^{M}\frac{s_i}{c_i} \quad (9)$$

For arbitrary positive number a, $\det(C) = a$, we can calculate the minimum value of $\mathrm{tr}(C^{-1}S)$ via Lagrange Multiplier. If $\mathrm{tr}(C^{-1}S)$ has a minimum value, then it means the value of (8) has an upper bound. The optimization problem here is

Minimize: $\mathrm{tr}(C^{-1}S)$

Subject to: $\det(C) = \prod_{i=1}^{M} c_i = a$

The Lagrange function is

$$F(\{c_i\}, \lambda) = \mathrm{tr}(C^{-1}S) + \lambda[\det(C) - a] \quad (10)$$

when $\begin{cases} c_i = \dfrac{s_i}{\lambda a} \\ \lambda = \dfrac{\det(S)^{\frac{1}{M}}}{a^{\frac{1}{M}}}\dfrac{1}{a} \end{cases}$, $\mathrm{tr}(C^{-1}S)$ will achieve its minimum value $\dfrac{\det(S)^{\frac{1}{M}}}{\det(C)^{\frac{1}{M}}}M$, and $L'(C;S) \le -\ln\det(C) - \dfrac{\det(S)^{\frac{1}{M}}}{\det(C)^{\frac{1}{M}}}M$, which means that $L'(C;S)$ has an upper bound when C is positive definite.

If C is singular, i.e., $\det(C) = 0$, the value of $L'(C;S) = -\ln\det(C) - \mathrm{tr}(C^{-1}S)$ tends to be minus infinity, which means $p(x_1, x_2, \ldots, x_N)$ tends to be zero, thus, trivial.

Therefore, we proved that $L'(C;S)$ has an upper bound when C is positive definite, and tends to be minus infinity when C is singular.

Form this theorem, we know that for any nonnegative definite matrix C, the value of $L'(C;S)$ will be less or equal to $-\ln\det(C) - \dfrac{\det(S)^{\frac{1}{M}}}{\det(C)^{\frac{1}{M}}}M$.

In the following part of this section, we will prove that the solution to the maximum-likelihood problem does exist. However, before starting proving, one definition and one lemma need to be introduced.

Definition: For a positive number e, there is a set of es $M_e = \{M \in R^{M\times M} \,|\, |m_{ij}| \le e\}$, where M is a $M \times M$ matrix, and m_{ij} is the entry lying in the i^{th} row and j^{th} column of the matrix.

Lemma 2: If a matrix A a real valued nonnegative definite symmetric matrix, a_{mn} is the largest element among all entries in A with respect to magnitude, then we have $\mathrm{tr}(A) \ge |a_{mn}|$. (The proof of this lemma can be found in Appendix.)

Now we start to prove the existence of the solution:

The set of nonnegative definite symmetric matrices is denoted as M_S. The intersection of M_e and M_S is shown in fig.3. It is clear that M_S is closed, and the boundary of M_S is the set of singular nonnegative definite symmetric matrices. The boundary of $M_e \cap M_S$ contains two parts: one is the set of singular nonnegative definite symmetric matrices (drawn in solid line), and the other is the set of positive definite symmetric matrices with one element on the main diagonal equal to e (drawn in the dash line). Moreover, from theorem 1, it is clear that the finite values of $L'(C;S)$ can only be obtained in the interior of M_S.

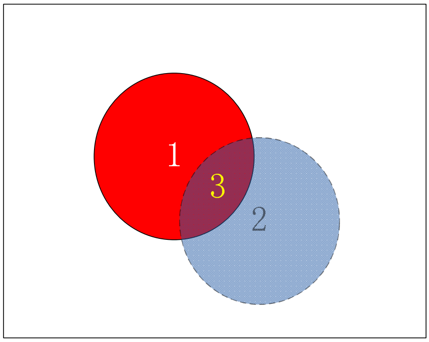

Fig. 3. The intersection of M_S and M_e. The area 1 denotes the set of M_S. The area 2 denotes the set of M_e, and the area 3 denotes the intersection of them.

To prove the existence of such a solution, we want to prove that for a large enough positive number e, the solution lies in the intersection of M_S and M_e. It is proved by contradiction. If such an e does not exist, it means for any e, there will always be an e' larger than e, and the optimal value of $L'(C;S)$ in $M_S \cap M_{e'}$ will be greater than that in$M_S \cap M_e$. Intuitively, the value of $L'(C;S)$ will be maximized when e goes to positive infinity. From lemma 1, we know there is a unitary matrix U which will transform C and S into diagonal matrices Λ_C and Λ_C simultaneously. Without losing generality, assume that c_k and s_l are the largest and smallest diagonal elements in Λ_C and Λ_C respectively. Form (9), $L'(C;S)$ can be cast as

$$L'(C;S) = -\sum_{i=1}^{M}(\ln c_i + \frac{s_i}{c_i}) = -\left(\ln c_k + \frac{s_k}{c_k}\right) - \sum_{i=1,i\neq k}^{M}(\ln c_i + \frac{s_i}{c_i})$$
$$\leq -\left(\ln c_k + \frac{s_l}{c_k}\right) - \sum_{i=1,i\neq k}^{M}(\ln c_i + \frac{s_i}{c_i}) \tag{11}$$

The inequality holds for c_k and s_l are the largest and smallest diagonal elements in Λ_C and Λ_C respectively. From lemma 2, if the largest element in C is c_{mn} which is equal to e, it is obvious that

$$Mc_k \geq tr(C) \geq c_{mn} = e \tag{12}$$

Therefore, $c_k \geq \frac{e}{M}$, which indicates that when e goes to positive infinity, c_k will also go to positive infinity. What is more, when c_k tends to positive infinity, the value of $L'(C; S)$ will be minus infinity. Hence, it contradicts the hypothesis that when e goes to positive infinity, $L'(C; S)$ will be maximized. So, the solution to the estimation problem must exist and can be found in the intersection of M_S and M_e with certain large e.

4 The Method of Solving the Estimation Problem

In this section, the method of solving the estimation problem is acquired through total differential of a function. Recall the objective function is

$$L'(C; S) = -\ln\det(C) - \mathrm{tr}(C^{-1}S)$$

when this function achieves its maximum, the total differential of this function will be zero.

Note that the objective function is the function with matrix parameter. The total differential of a $F(M)$, where M is a matrix parameterized by $m_1, m_2, \dots, m_k$, is

$$d\big(F(M)\big) = \sum_{i=1}^{k} \frac{\partial (F(M))}{\partial m_i} dm_i \tag{13}$$

For example, if $F(M) = M$, and M is parameterized by all its elements, then the total differential of $F(M)$ is

$$dF(M) = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} dm_{11} + \cdots + \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} dm_{NN} = \begin{bmatrix} dm_{11} & \cdots & dm_{1N} \\ \vdots & \ddots & \vdots \\ dm_{N1} & \cdots & dm_{NN} \end{bmatrix}$$

The total differential of the first part of (8) is

$$d(\det(C)) = \sum_{i=1}^{M} \frac{\partial \sum_{j=1}^{M} c_{ij} C_{ij}^*}{\partial c_{ij}} dc_{ij} = \mathrm{tr}(C^{-1} dC) \tag{14}$$

where the C_{ij}^* denotes the adjoint matrix of c_{ij}.

The total differential of the second part of (8) is

$$d\big(\mathrm{tr}(C^{-1}S)\big) = -\mathrm{tr}(C^{-1} dC C^{-1} S) \tag{15}$$

Therefore, the total differential of (8) is

$$d\big(L'(C; S)\big) = \mathrm{tr}(C^{-1} dC C^{-1} S - C^{-1} dC) = \mathrm{tr}[(C^{-1} S C^{-1} - C^{-1}) dC] \tag{16}$$

For any feasible direction of variation of C, the total differential must be zero, i.e., $\mathrm{tr}[(C^{-1}SC^{-1} - C^{-1})dC]$ must be zero. Especially, when dC meets the structure requirements of C, and since (16) is always zero for arbitrary direction of variation of C, we can substitute dC by C, which leads to

$$\mathrm{tr}(C^{-1}S) = M \tag{17}$$

From this, it is clear that if there are no further constraints on C, the best estimator of C is S itself.

Till now, only the real valued covariance matrices estimation is taken into consideration. However, from the form of array manifold matrix, we know that the matrices of interest are complex valued matrices. Nevertheless, we can alter the estimation problem of complex valued matrices to the estimation of real valued matrices via reconstructing S.

Recall that in the real valued scenario, $S=\frac{1}{N}\sum_{i=1}^{N} x_i x_i^T$. Similarly, when the elements of x_i are complex valued, $S=\frac{1}{N}\sum_{i=1}^{N} x_i x_i^H$. We can struct $S^R=\frac{1}{N}\sum_{i=1}^{N}\begin{bmatrix}Re(x_i)\\ Im(x_i)\end{bmatrix}\begin{bmatrix}Re(x_i)^T & Im(x_i)^T\end{bmatrix}$ which is obviously real valued, and positive definite symmetric as well. We can use the real valued and positive definite symmetric matrix S^R to estimate C^R.

Let $O=\frac{1}{N}\sum_{i=1}^{N} Re(x_i)Re(x_i)^T$, $P=\frac{1}{N}\sum_{i=1}^{N} Im(x_i)Im(x_i)^T$, and $Q=\frac{1}{N}\sum_{i=1}^{N} Re(x_i)Im(x_i)^T$, thus, $S^R=\begin{bmatrix}O & Q\\ Q^T & P\end{bmatrix}$. The estimation of C is acquired by averaging the outer product of N sensed data. What is more, the relationship between C^R and C is one-to-one mapping, i.e., $c_{ij}=o_{ij}+p_{ij}-jq_{ij}+jq_{ji}$.

Therefore, we can acquire the estimation of a complex valued matrix via estimation of a corresponding real values matrix. It is shown in the following example that how to acquire the arrival angels.

Suppose there are N events of interest and M sensor in the field. $x_1, x_2, \ldots, x_N$ are N samples we acquired, and each sample $x_i$is a M-dimension vector with each element equal to the signal of the i^{th} event acquired by a sensor. Each vector is complex valued. The matrix $S=\frac{1}{N}\sum_{i=1}^{N} x_i x_i^H$ is used to estimate the covariance matrix C.

From above discussion the estimator of C can be obtained from the estimation of a real valued matrix $S^R=\begin{bmatrix}O & Q\\ Q^T & P\end{bmatrix}$ via Maximum-likelihood approach. Denote C' as the estimator of C. From section 2, the covariance matrix has the form of (1). Suppose that the received signals are interfered by some additive noise with zero mean and σ_n^2 variance Gaussian distribution, and the i^{th} power of received signals is σ_i^2. We can acquire the array manifold matrix G through the following equation:

$$C' = \sigma_n^2 I_M + \sum_{i=1}^{N} \sigma_i^2 g_i g_i^H \tag{19}$$

Every g_i is with the form $\begin{bmatrix}1 & \exp\left(-\frac{j2\pi d}{\lambda}\sin\theta_1\right) & \ldots & \exp\left(-\frac{j(M-1)2\pi d}{\lambda}\sin\theta_N\right)\end{bmatrix}$, which means the each arrival angle θ_i is now available to us.

5 Conclusion

In this article, intensive study has been made in order to estimate the angles of arrival of signals sensed by sensor nodes in linear wireless sensor networks. In dealing with this problem, we use signal covariance matrix estimation techniques after the discussion of the relationship between arrival angles and signal covariance matrices. The maximum-likelihood estimation approach is adopted to solve the estimation problem. In future researches, the relationship between arrival angles and covariance matrices needs to be further studied, which will reveal a much more explicit insight into the relationship between these subjects. Studies of other methods, such as parameter estimation, which will give us a direct way to acquire the arrival angles, and minimum entropy estimation, will also be carried out successively.

References

1. Zimmerling, M., Dargie, W., Reason, J.M.: Energy-Efficient Routing in Linear Wireless Sensor Networks. In: 2007 IEEE International Conference on Mobile Ad-hoc and Sensor Systems, vol. 1-3, pp. 210–212 (2007)
2. Gibson, J., Xie, G.G., Yang, X.: Performance Limits of Fair-Access in Sensor Networks with Linear and Selected Grid Topologies. In: GLOBECOM 2007: 2007 IEEE Global Telecommunications Conference, vol. 1-11, pp. 688–693 (2007)
3. Noori, M., Ardakani, M.: Characterizing the traffic distribution in linear wireless sensor networks. IEEE Communication Letters 12(8), 554–556 (2008)
4. Hong, L., Xu, S.: Energy-Efficient Node Placement in Linear Wireless Sensor Networks. In: International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), vol. 2, pp. 104–107 (2010)
5. Ahmed, A.A., Shi, H., Shang, Y.: A survey on network protocols for wireless sensor networks. In: Proceedings of the International Conference on Information Technology: Research and Education, pp. 301–305 (August 2003)
6. Ye, W., Heidemann, J., Estrin, D.: Medium access control with coordinated adaptive sleeping for wireless sensor networks. IEEE/ACM Transactions on Networking 12(3), 493–506 (2004)
7. Werner, K.: Kronecker: structured covariance matrix estimation. In: Proceedings of 2007 IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 3, pts. 1-3, pp. 825–828 (2007)
8. Richter, A.: ML estimation of covariance matrix for tensor valued signals in noise. In: 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 1-12, pp. 2349–2352 (2008)
9. Jansson, M.: ML Estimation of Covariance Matrices with Kroneckor and Persymmetirc Strucure. In: Proceedings of 2009 IEEE 13th Digital Signal Processing Workshop & 5th IEEE Proessing Education Workshop, vol. 1, 2, pp. 298–301 (2009)
10. Werner, K.: On Estimation of Covariance Matrices With Kronecker Product Structure. IEEE Transactions on Signal Processing 56(2), 478–491 (2008)

Appendix

Toeplitz Matrix

A matrix T is a toeplitz matrix if arbitrary element t_{ij} in T is equal to t_{i-j}, i.e.,

$$T = \begin{bmatrix} t_0 & \cdots & t_{1-M} \\ t_1 & \cdots & t_{2-M} \\ \cdots & \cdots & \cdots \\ t_{M-1} & \cdots & t_0 \end{bmatrix}$$

$$\sum_{i=1}^{N} x_i^T C^{-1} x_i = \sum_{i=1}^{N} tr(x_i^T C^{-1} x_i) = \sum_{i=1}^{N} tr(C^{-1} x_i x_i^T) = \sum_{i=1}^{N} C^{-1} tr(x_i x_i^T)$$

Lemma 1

Proof: if A and B are positive definite symmetric, the matrix A+B is also positive definite symmetric. It is obvious that a positive definite symmetric matrix is congruent to identity matrix I. Therefore, A+B is congruent to I, and the congruence matrix is U1, i.e., $U_1^T(A + B)U_1 = I$. It is clear that the matrix $U_1^T A U_1$ is positive definite symmetric as well, hence, there exists one unitary matrix U_2 which will change $U_1^T A U_1$ into a diagonal matrix Λ_A, i.e., $U_2^T U_1^T A U_1 U_2 = \Lambda_A$. Hence,

$$U_2^T U_1^T (A + B) U_1 U_2 = \Lambda_A + U_2^T U_1^T B U_1 U_2 = I.$$

Therefore, $U_2^T U_1^T B U_1 U_2 = I - U_2^T U_1^T A U_1 U_2$, which means that $U_2^T U_1^T B U_1 U_2$ is also a diagonal matrix. So, there exists one unitary matrix $U_1 U_2$ which can shape A and B into diagonal matrix simultaneously.

Lemma 2

Proof: if a_{mn} is positive, we construct a vector with mth and nth entries equal to 1 and -1. Then we have

$$x^T A x = a_{mm} + a_{nn} - a_{mn} - a_{nm} = a_{mm} + a_{nn} - 2a_{mn} \geq 0$$

Because A is nonnegative definite, it elements on the main diagonal are all greater than or equal to zero. Therefore, we have $tr(A) \geq 2a_{mn} \geq a_{mn}$.

If a_{mn} is negative, we construct a vector with mth and nth entries both equal to 1. Then we have

$$x^T A x = a_{mm} + a_{nn} + a_{mn} + a_{nm} = a_{mm} + a_{nn} + 2a_{mn} \geq 0$$

Therefore, we have $tr(A) \geq -2a_{mn} \geq -a_{mn} = |a_{mn}|$.