

样本协方差矩阵的特征值的中偏差原理

金鑫, 解永晓*

山东师范大学数学与统计学院, 山东 济南

Email: nextstop_happiness@163.com, *yxx8511@163.com

收稿日期: 2021年3月27日; 录用日期: 2021年4月15日; 发布日期: 2021年4月30日

摘要

样本协方差矩阵特征值的大偏差对移动通信领域中比特错误问题的近似计算有非常重要的应用。设矩阵 $C_{k \times n}$ 的元素 C_{ij} 独立且满足 $E[C_{ij}] = 0$, $Var(C_{ij}) = 1$, 则样本协方差矩阵 $W = \frac{1}{n}CC^T$ 的特征值为 $\frac{1}{n} \sum_{i=1}^n S_{x,i}^2 = \frac{1}{n} \sum_{i=1}^n \left(\sum_{m=1}^k x_m C_{mi} \right)^2$, 其中 $x = (x_1, x_2, \dots, x_k)$, 满足 $\|x\|_2 = 1$ 。本文首先给出样本协方差矩阵特征值的中偏差原理, 即 $\{S_{x,i}^2\}$ 满足速度为 $\frac{n}{B^2(n)}$ 速率函数为 $I(x) = \frac{x^2}{2\sigma^2}$ 的中偏差原理, 其中 $\frac{B(n)}{\sqrt{n}} \rightarrow \infty$, $\frac{B(n)}{n} \rightarrow 0$, $(n \rightarrow \infty)$, $Var(S_{x,i}^2) = \sigma^2$ 。其次使用数值模拟的方法利用python验证了该定理的准确性, 最后利用该中偏差原理来对比CDMA系统中不同解码方案的优劣。

关键词

样本协方差矩阵, 中偏差原理, 特征值, 码分多址

Moderate Deviation Principle of the Eigenvalues of the Sample Covariance Matrix

Xin Jin, Yongxiao Xie*

School of Mathematics and Statistics, Shandong Normal University, Jinan Shandong

Email: nextstop_happiness@163.com, *yxx8511@163.com

Received: Mar. 27th, 2021; accepted: Apr. 15th, 2021; published: Apr. 30th, 2021

*通讯作者。

Abstract

The large deviation of eigenvalues of sample covariance matrix has great significance in calculating the bit error probability of Code Division Multiple Access. Throughout the paper, we assume that the i.i.d. real matrix elements of $C_{k \times n}$ satisfy $E[C_{ij}] = 0$, $Var(C_{ij}) = 1$, then the eigenvalue of the sample covariance matrix $W = \frac{1}{n}CC^T$ is $\frac{1}{n} \sum_{i=1}^n S_{x,i}^2 = \frac{1}{n} \sum_{i=1}^n \left(\sum_{m=1}^k x_m C_{mi} \right)^2$, where x is with k coordinates and norm $\|x\|_2 = 1$. First of all, we prove that $\{S_{x,i}^2\}$ satisfies moderate deviations where speed is $\frac{n}{B^2(n)}$ and the rate function is $I(x) = \frac{x^2}{2\sigma^2}$, where $\frac{B(n)}{\sqrt{n}} \rightarrow \infty$, $\frac{B(n)}{n} \rightarrow 0$, ($n \rightarrow \infty$), $Var(S_{x,i}^2) = \sigma^2$. Then we use the numerical simulation method to model the theoretical function by python, to verify its reliability. Finally, we use the moderate deviation of sample covariance matrix in the CDMA to compare the advantages and disadvantages of different decoding techniques.

Keywords

Sample Covariance Matrix, Moderate Deviation Principle, Eigenvalues, Code Division Multiple Access

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

中偏差(Moderate Deviation Principle)及大偏差(Large Deviation Principle)原理可以度量稀有事件概率的渐近性质,特别是独立同分布的随机变量序列的中偏差及大偏差原理。因其典型的意义,在物理、金融和通信等很多方面都有显著的应用。虽然独立同分布随机变量序列的中偏差原理和大偏差原理形式相近,但在一定程度上还是有本质区别的[1]。

样本协方差矩阵在统计学中经常被推广到各类实际问题,如在电力、通信、金融和经济中进行建模。样本协方差矩阵的特征值是其中的一个重要元素,在实际问题中可以表示出很多具有现实意义的数。考虑一个 n 行 k 列的矩阵 $C_{n \times k}$, 它的样本协方差矩阵为 $W = \frac{1}{n}CC^T$, 则对于任意的向量 $x = (x_1, x_2, \dots, x_k)$,

满足 $\|x\|_2 = 1$ 。我们就可以用 $\langle x, Wx \rangle$ 表示 W 的特征值,而这个 $\langle x, Wx \rangle$ 有实际的意义。比如 $C_{n \times k}$ 表示 CDMA 系统中的码相关矩阵,则借助 $\langle x, Wx \rangle$ 的取值概率可以表示判断解码方案优劣的一种指标——比特错误概率。

Bai 等人[2] [3] [4] [5]分别研究了元素为独立同分布随机变量时样本协方差矩阵的最大特征值和最小特征值的渐近行为。特别地,当 C 中的元素服从标准正态分布时, W 被称作 Wishart 矩阵, Wishart 矩阵在统计学中扮演着重要的角色,其样本特征值的大偏差原理可见参考文献[6]。Guionnet [7]、Hiai 和 Petz [8]研究了速度为 $\frac{1}{n^2}$ 时 Wishart 矩阵的样本特征值的大偏差原理。Anne 等人[6] [9]证明了样本协方差

矩阵最大最小特征值的大偏差原理, 并在[3]中表明: 当 $n \rightarrow \infty$, 且 k 固定或者趋于无穷但不快于 $o(n/\log \log n)$ 时, W 的特征值是趋向于 1 的。由于 W 的最大最小特征值大偏差恰好就是 BEP (Bit Error Probability), 最后 Anne 等人将样本协方差矩阵的特征值大偏差情况的结果应用到 CDMA 系统的不同解码方案的建模中, 给出了衡量 CDMA 解码系统优劣的另一种标准, 该方法优于之前所使用的 SNR (Signal Noise Ratio), 见参考文献[10]。但样本协方差矩阵的最大最小特征值的大偏差速率函数没有办法直接求解, 所以 Anne 等人在 Wishart 矩阵中通过中心极限定理逼近的方法得到所求的样本协方差矩阵特征值的速率函数的渐近速率函数, 并且在 $\lambda_{\min} = 0$ 问题没有得到很好的解决, 于是本文通过使用大偏差原理的普适性来尝试辅助解决这一问题。

定义 1.1 考虑一系列独立同分布的随机变量序列 $\{Y_n\}$, 称 $\mathbb{P}(Y_n \in \cdot)$ 在拓扑空间 X 上服从速度为 a_n 且速率函数为 $I(\cdot): X \rightarrow R^+$ 的大偏差原理, 是指

1) 对所有的开集 $G \subseteq X$, 有

$$\liminf_{n \rightarrow \infty} \frac{1}{a_n} \log \mathbb{P}(Y_n \in G) \geq -\inf_{x \in G} I(x), \tag{1}$$

2) 对所有的闭集 $A \subseteq X$, 有

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} \log \mathbb{P}(Y_n \in A) \leq -\inf_{x \in A} I(x), \tag{2}$$

其中 I 是下半连续函数, 并且对所有的 $L \in R^+$, 有紧的水平集 $N_L := \{x \in X : I(x) \leq L\}$ 。

若 $\{Y_n\}$ 在拓扑空间 X 上服从速度为 $\frac{n}{B^2(n)}$ 且速率函数为 $I(\cdot): X \rightarrow R^+$ 的中偏差原理, 其中 $B(n)$ 是单调递增的数列, 并且当 $n \rightarrow \infty$ 时满足

$$\frac{B(n)}{n} \rightarrow 0, \quad \frac{B(n)}{\sqrt{n}} \rightarrow \infty, \tag{3}$$

在形式上也有类似的结果。

定义 1.2 对于如定义 1.1 中的 $\{Y_n\}$, 记 $S_n = \sum_{i=1}^n Y_n$, 称 $\mathbb{P}\left(\frac{S_n}{B(n)} \in \cdot\right)$ 满足速度为 $\frac{n}{B^2(n)}$, 速率函数为 $I(x)$ 的中偏差原理, 是指

1) 对所有的开集 $G \subseteq X$, 有

$$\liminf_{n \rightarrow \infty} \frac{n}{B^2(n)} \log \mathbb{P}\left(\frac{S_n}{B(n)} \in G\right) \geq \inf_{x \in G} I(x)$$

2) 对所有的闭集 $A \subseteq X$, 有

$$\limsup_{n \rightarrow \infty} \frac{n}{B^2(n)} \log \mathbb{P}\left(\frac{S_n}{B(n)} \in A\right) \leq \inf_{x \in A} I(x)$$

其中 $I(x) = \frac{x^2}{2\text{Var}(Y_1)}$ 。

大偏差原理和中偏差原理有本质的区别。大偏差原理描述的是依赖于大数定律的遍历现象, 而中偏差原理则是描述的介于大数定律和中心极限定理之间的概率的渐近行为。具体地, 大偏差原理的速率函数的计算依赖于随机变量的分布, 而中偏差原理的速率函数只依赖于变量的中心极限定理的极限密度形

式.也就是说, 中偏差原理的速率函数并不依赖于随机变量的分布, 所以中偏差原理相比于大偏差原理有普适性的特点.

对于一系列独立同分布随机变量序列 $X_1, X_2, \dots, X_n \dots$ 若

$$\exists \delta > 0, Ee^{\delta|X_1|} < +\infty, \quad (4)$$

则它的对数矩生成函数 $\Lambda(t) = \log \mathbb{E}[e^{tX_1}] < \infty, \forall t < \infty$, 由Cremér定理可知, 随机变量序列 $\{X_n\}$ 的部分和序列 $\{S_n\}$ 满足速度为 $\frac{1}{n}$, 速率函数为 $I(\alpha)$ 的大偏差原理, 其中速率函数 $I(\alpha) = \sup_t (t\alpha - \log \mathbb{E}[e^{tX_1}])$. 根据参考文献[11]的命题 1.7 可得, 在(4)的条件下, 对于满足条件(3)的 $B(n)$, 随机变量序列 $\{X_n\}$ 满足速度为 $\frac{n}{B^2(n)}$ 速率为 $I(\cdot)$ 的中偏差原理, 其中

$$I(x) = \frac{x^2}{2\text{Var}(X_1)}. \quad (5)$$

2. 样本协方差矩阵的特征值的中偏差原理

2.1. 系统模型

对于样本协方差矩阵为 $W = \frac{1}{n}CC^T$, 如果矩阵 $C_{k \times n}$ 中的元素含有随机变量, 则 W 和 C 都是随机矩阵. 本文假设 C 中的元素都是独立同分布的随机变量且满足

$$\mathbb{E}(C_{ij}) = 0, \text{Var}(C_{ij}) = 1. \quad (6)$$

对于满足上述条件的 C , 当元素 C_{ij} 的四阶矩有限, 并且 $k, n \rightarrow \infty$, $\frac{k}{n} = \beta$ 时, 其中 β 是一常数, 那么 W 的特征值会收敛到固定的值. 进一步, 根据 Bai 和 Yin 的文献[5]可知, λ_{\max} 收敛到 $(1 + \sqrt{\beta})^2$, 而 λ_{\min} 收敛到 $(1 - \sqrt{\beta})^2$, 其中 $x_+ = \max\{0, x\}$. 对于 C_{ij} , 因为满足条件(6), 可得 W 的非对角元素趋于 0, 而对角元素趋于 1. 对于矩阵 W 和任意的满足 $\|x\|_2 = 1$ 的 k 维向量 x , 可知

$$\lambda_{\min} \leq \langle x, Wx \rangle \leq \lambda_{\max},$$

其中 $\langle x, Wx \rangle$ 可以表示为

$$\langle x, Wx \rangle = \frac{1}{n} |C^T x|^2 = \frac{1}{n} \sum_{i=1}^n \left(\sum_{m=1}^k x_m C_{mi} \right)^2.$$

记

$$S_{x,i} = \sum_{m=1}^k x_m C_{mi}.$$

根据 C 的定义, 可得 $\{S_{x,i}\}$ 是独立同分布的随机变量序列, 并且满足 $\mathbb{E}(S_{x,i}) = 0, \text{Var}(S_{x,i}) = \|x\|^2$, 通过Cremér定理可以得到 $\{S_{x,i}^2\}$ 的大偏差原理, 具体的内容可以查阅参考文献[6]和[9].

本文主要考虑的是 $\{S_{x,i}^2\}$ 的中偏差定理, 下面先来看一下主要结果.

2.2. 主要结果

对于上述定义的序列 $\{S_{x,i}\}$, 记 $\sigma^2 = \text{Var}(S_{x,i}^2)$, $S_n = \sum_{i=1}^n S_{x,i}^2$. 设 $B(n)$ 为单调上升序列, 满足(3)式, 若 $\exists \epsilon > 0$, 使得

$$\mathbb{E}\left[e^{\epsilon C_1^2}\right] < \infty \tag{7}$$

则随机变量序列 $\{S_{x,i}^2\}$ 满足中偏差原理, 下面给出 $S_{x,i}^2$ 的具体中偏差结果。

定理 2.2.1 记 $\mu_n = \mathbb{P}\left(\left|\frac{S_n - n}{B(n)}\right| < \epsilon\right)$, 若(7)成立, 则 $\{S_{x,i}^2\}$ 满足速度为 $\frac{n}{B^2(n)}$, 速率为 $I(x)$ 的中偏差原理, 其中 $I(x) = \frac{x^2}{2\sigma^2}$, 即

$$\lim_{n \rightarrow \infty} \frac{n}{B^2(n)} \log \mathbb{P}\left(\left|\frac{S_n - n}{B(n)}\right| < \epsilon\right) = \frac{\epsilon^2}{2\sigma^2}. \tag{8}$$

证明: 由参考文献[11]知, $(\mu_n, n \rightarrow \infty)$ 的 Cremér 泛函为

$$\begin{aligned} \mathbb{P}(y) &= \lim_{n \rightarrow \infty} \log \int_{\mathbb{R}} \exp\langle x, y \rangle \frac{n}{B^2(n)} d\mu_n(x) \\ &= \lim_{n \rightarrow \infty} \frac{n}{B^2(n)} \log \mathbb{E} \exp\left[\frac{(S_n - n)y}{B(n)} \cdot \frac{B^2(n)}{n}\right] \\ &= \lim_{n \rightarrow \infty} \frac{n}{B^2(n)} \log \mathbb{E} \exp\left[(S_{x,1}^2 - 1) \cdot \frac{B(n)}{n}\right]^n \\ &= \lim_{n \rightarrow \infty} \frac{n}{B^2(n)} \log \mathbb{E} \exp\left[S_{x,1}^2 \cdot y \cdot \frac{B(n)}{n} - y \cdot \frac{B(n)}{n}\right] \\ &= \lim_{n \rightarrow \infty} \frac{n^2}{B^2(n)} \left[\Lambda\left(\frac{B(n)}{n} \cdot y\right) - \frac{B(n)}{n} \cdot y \right] \end{aligned}$$

其中

$$\Lambda(y) = \log \mathbb{E} \exp(S_{x,1}^2 y),$$

设 $f(t) = \Lambda(ty) = \log \mathbb{E} \exp(S_{x,1}^2 ty)$, 计算可以得

$$f'(t) = \frac{\mathbb{E} S_{x,1}^2 y \cdot e^{S_{x,1}^2 ty}}{\mathbb{E} e^{S_{x,1}^2 ty}},$$

当 $t=0$ 可得 $f'(0) = y$, 再对上式求导可得

$$f''(t) = \frac{\mathbb{E}(S_{x,1}^2 y)^2 \cdot e^{S_{x,1}^2 ty} - (\mathbb{E} S_{x,1}^2 y \cdot e^{S_{x,1}^2 ty})^2}{(\mathbb{E} \exp S_{x,1}^2 ty)^2},$$

带入 $t=0$ 可得 $f''(0) = \mathbb{E}(S_{x,1}^2 y)^2 - \mathbb{E}^2(S_{x,1}^2 y) = y^2 \cdot \sigma^2$ 。

由泰勒公式就可以得

$$\mathbb{P}(y) = \lim_{n \rightarrow \infty} \frac{1}{\left(\frac{B^2(n)}{n}\right)^2} \left[f\left(\frac{B(n)}{n}\right) - f'(0) \frac{B(n)}{n} \right],$$

所以

$$\mathbb{P}(y) = \frac{1}{2} f''(0) = \frac{y^2 \sigma^2}{2},$$

则它的 Legendre 变换即速率函数就是

$$I(x) = \sup \left(\langle x, y \rangle - \frac{\sigma^2 y^2}{2} \right) = \frac{x^2}{2\sigma^2}.$$

则

$$\lim_{n \rightarrow \infty} \frac{n}{B^2(n)} \log \mathbb{P} \left(\left| \frac{S_n - n}{B(n)} \right| < \varepsilon \right) = \frac{\varepsilon^2}{2\sigma^2}.$$

从而 $\mu_n = \mathbb{P} \left(\left| \frac{S_n - n}{B(n)} \right| < \varepsilon \right)$ 满足速率函数为 $I(x)$, 速度为 $\frac{n}{B^2(n)}$ 的大偏差原理, 即(8)式成立, 则随机变量序列 $\{S_{x,i}^2\}$ 的中偏差定理得证。

2.3. 数值模拟

对于所求速率函数 $I(x)$, 因为并没有良好的验证方法去验证其有效性, 于是本文考虑(8)式, 通过数值模拟的方法去实现对中偏差估计的检验。

上述中偏差的具体形式为定理 2.2.1 中的(8)式, 当 n 充分大时, 由(8)式得

$$-\frac{n}{B^2(n)} \log \mathbb{P} \left(\left| \frac{S_n - n}{B(n)} \right| < \varepsilon \right) \sim \frac{\varepsilon^2}{2\sigma^2}$$

从而

$$\mathbb{P} \left(\left| \frac{S_n - n}{B(n)} \right| < \varepsilon \right) \sim e^{-\frac{(\varepsilon B(n))^2}{2n\sigma^2}},$$

记

$$P(n) = \mathbb{P} \left(\left| \frac{S_n - n}{B(n)} \right| < \varepsilon \right),$$

$$Q(n) = e^{-\frac{(\varepsilon B(n))^2}{2n\sigma^2}}.$$

通过借助 python 生成 n 个满足 $\mathbb{E}[x_n] = 0, \text{Var}(x_n) = 1$ 的随机变量 x_n , 其中 $x_n, n = 1, 2, \dots, 5000$, 然后通过带入 x_n 计算 $P(n) = \mathbb{P} \left(\left| \frac{S_n - n}{B(n)} \right| < \varepsilon \right)$ 的图像, 并可以直接输入公式 $Q(n) = e^{-\frac{(\varepsilon B(n))^2}{2n\sigma^2}}$ 来画出 $P(n)$ 的图像。本例中取 $B(n) = n^{0.9}$, $\varepsilon = 0.1$, 最后将这两个函数的图像放在同一坐标系中来判断拟合的准确程度, 最终结果如下图。

图 1 中, 橙色曲线为 $P(n)$ 的图像, 蓝色曲线为 $Q(n)$ 的图像, 根据图片可以看出, 随着 n 的逐渐增大, $P(n)$ 逐渐逼近 $Q(n)$ 。具体地: 当 $n < 500$ 时, 两函数之间还会有些差异, 这应该是单次模拟不充分导致的, 但是当 $n \in (500, 1000)$ 时, $P(n)$ 就渐渐地逼近 $Q(n)$, 至于 $n > 1000$ 时, 拟合的结果就非常好了, 这是符合定理 2.2.1 的预期结果。

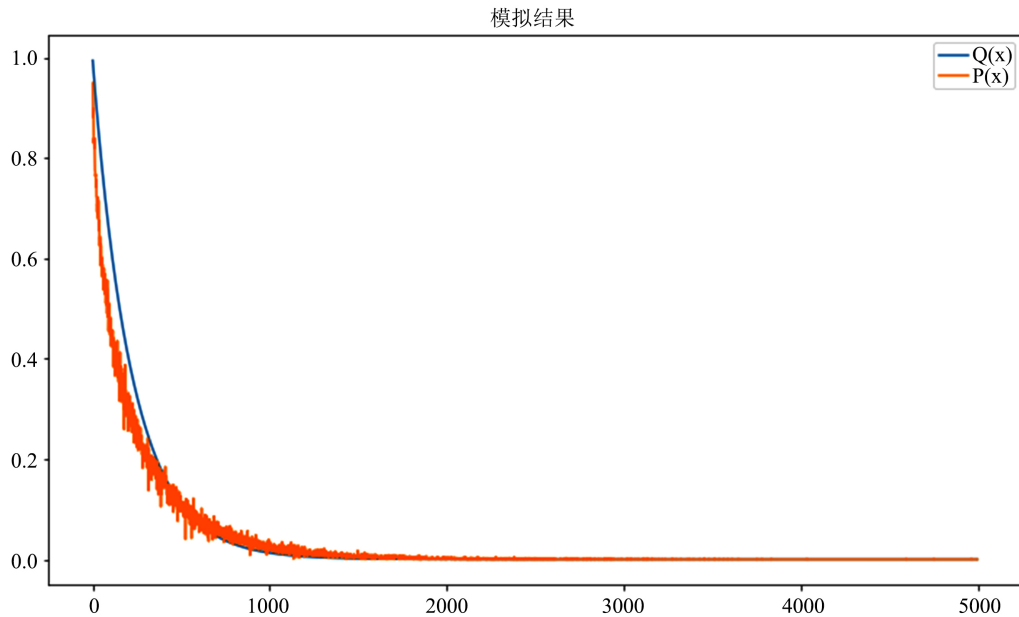


Figure 1. Simulation result of Theorem 2.2.1
图 1. 定理 2.2.1 的模拟结果

3. 在 CDMA 系统中的应用

下面先简单地介绍一下本文所用到的 CDMA 系统。考虑共有 k 个用户的 CDMA 系统，每个用户发送的都是长度为 n 的信号，且每个信号中的元素都满足 $\mathbb{P}(C_{mi} = -1) = \mathbb{P}(C_{mi} = 1) = \frac{1}{2}$ ，第 m 个用户所发送的信号为 $C_m = (c_{m1}, c_{m2}, \dots, c_{mn})$ ，整个系统中的用户所发送的信号为一个 k 行 n 列的矩阵 C ，所以基站所接受的总体信号 s 就可以概括成

$$s = \sum_{m=1}^k r_m + N = C^T b + N,$$

其中 N 是由 n 个独立的高斯随机变量组成， $N_i \sim N(0, \sigma^2)$ ， $i = 1, 2, \dots, n$ ， σ 表示噪声强弱的参数，即可加高斯白噪声(AWGN)。因为组成 s 的向量会在同一时间受到同一噪音的影响，并且有很多方法可以消除掉可加高斯白噪声，所以它对整个模型的影响是可以控制的，通常都会选择固定参数 σ 进行建模，具体可见参考文献[6] [9]。

当基站接收到用户发来的编码后，需要对码序列进行解码，比如对用户 m 进行解码，就用码相关矩阵 C 和用户 m 的码序列 c_m 做内积，然后再除以 n 。此时，得到 b_m 的一个估计值 \hat{b}_m ，

$$\hat{b} = \frac{1}{n} C C^T b + \frac{1}{n} C N.$$

由于矩阵 $\frac{1}{n} C C^T$ 是样本协方差矩阵 W ，上式变成了

$$\hat{b} = \frac{1}{n} W b + \frac{1}{n} C N,$$

通过变形可以把 \hat{b}_m 写成

$$\hat{b} = b + (W - I) b + \frac{1}{n} C N,$$

由这个式子可以清楚地看出, \hat{b}_m 是由真实值 b 、多址干扰 $(W-I)b$ 和 $\frac{1}{n}CN$ (即 AWGN) 组成。其中多址干扰(MAI)是影响解码系统效率的一个重要因素, 在判断解码系统是否出现了问题时, 主要是判断解码后得到的用户码序列的估计值 \hat{b}_m 和原本的用户码序列 b_m 的符号是否一致。当 $\hat{b}_m = 0$ 时符号就需要等可能的随机来判定正负, 而如果发现 \hat{b}_m 和 b_m 的符号不一致时, 就认为解码系统产生了一个比特错误, 这就是匹配滤波器解码(Matched Filter)出现错误的原理。

相对 MF 解码来说更加有优势的解码系统就是 PIC 解码(Parallel Interference Cancellation), 有 SD-PIC (Soft-Decision Parallel Interference Cancellation)和 HD-PIC (Hard-Decision Parallel Interference Cancellation) 解码方案。在大多数的解码方案中习惯性地多址干扰和可加高斯白噪声放到一起进行处理。这种处理方式会影响对 CDMA 系统总体容量的估计, 从而产生较大的误差, 但是 PIC 解码会单独处理 MAI, 就会避免这种问题。

在多阶 SD-PIC 方案中, 考虑忽略噪音的情况。若用 λ_j 表示 W 的第 j 个特征值, 一组正交基 w_j 是与之对应的特征向量, 向量 b 可用 W 的特征向量 w_j 来表示。

$$b = \sum_{j=1}^k \beta_j w_j,$$

则 SD-PIC 方案的第 s 步迭代 $\hat{b}^{(s)}$ 可用 λ_j 表示,

$$\hat{b}^{(s)} = \sum_{\zeta=0}^{s-1} \sum_{j=1}^k (1-\lambda_j)^\zeta \lambda_j \beta_j w_j = \sum_{j=1}^k \left(1 - (1-\lambda_j)^s\right) \beta_j w_j,$$

在忽视噪音的情况下, 若出现特征值等于 0 或大于 2 的情况, 就会导致 $\left(1 - (1-\lambda_j)^s\right)$ 不收敛, 从而产生比特错误。所以样本协方差矩阵 W 的最大、最小特征值的取值决定了是否会发生比特错误, 我们可以通过计算样本协方差矩阵的最大、最小特征值偏差的概率来描述比特错误发生的概率, 这样就把 CDMA 系统解码错误优劣的比较问题抽象成了一个数学问题。参考文献[2][3]详细地介绍了整个建模过程, 而关于 CDMA 系统的具体知识, 可以阅读[7]。因为发生比特错误是一个稀有事件, 所以 BEP 可以通过使用大偏差原理来求解, 在文献[3]中通过近似 Wishart 矩阵的方式计算出了该系统的 BEP 的速率函数, 并分别给出了 SD-PIC 和 dec 两种解码方案的具体的 BEP 的速率函数, 这里设该解码系统的大偏差原理速率函数为 $I'(x)$, 对于 SD-PIC, 有 $I'_{SD} \geq \min\{I'_{\max}(2), I'_{\min}(0)\}$ 。对于 dec, 有 $I'_{dec} \geq I'_{\min}(0)$, 所以通过计算可以给出具体数值

$$I'_{SD} \geq \frac{1}{2} - \frac{1}{2} \log 2,$$

$$I'_{dec} \geq \log 2,$$

根据大偏差原理 $BEP \leq e^{-nI}$ 就可以计算给定用户数 k 和编码长度 n 的 CDMA 系统的不同解码方案的比特错误概率来比较它们的优劣。

在上述的 CDMA 系统中, 由于 $\mathbb{P}(C_{mi} = -1) = \mathbb{P}(C_{mi} = 1) = \frac{1}{2}$ 是满足条件(6)的, 所以可利用本文的结果来求解该 CDMA 系统的中偏差速率函数 $I(n)$, 所以在中偏差情况下, 对于 SD-PIC, 相应的中偏差速率函数就是

$$I_{SD} \geq \min \left\{ I \left(\frac{0-n}{B_n} \right), I \left(\frac{2-n}{B_n} \right) \right\},$$

对于 dec 解码相应的中偏差速率函数就是

$$I_{dec} \geq I\left(\frac{0-n}{B_n}\right),$$

对于具体情况, 比如考察 $x = \frac{1}{\sqrt{k}}(1, 1, \dots)$ 的情况, 此时 $\sigma^2 = 1$, 那中偏差速率函数 $I(\alpha) = \frac{\alpha^2}{2}$, 于是对于 $n = 16$, $\lambda \leq 2$ 的速率函数即可求得 $I(\lambda \leq 2) \approx 0.405$, 这与 $I'_{dec} \geq \frac{1}{2} \log 2 \approx 0.346$ 的结果相近, 再将 $I(\lambda \leq 2)$ 带入 $Q(x)$ 即可算出 $P(dec)$ 得到最终结果

$$P(dec) \approx 0.024,$$

也就是说当一个 CDMA 系统中有 16 个用户时, 使用 dec 解码方案会出错的概率约为 0.024。

基金项目

国家自然科学基金青年基金(批准号: 11601287)。

参考文献

- [1] Arcones, M.A. (2003) Moderate Deviations of Empirical Processes. *Stochastic Inequalities and Applications*, **56**, 189-212. https://doi.org/10.1007/978-3-0348-8069-5_13
- [2] Bai, Z.D. and Silverstein, J.W. (1998) No Eigenvalues outside the Support of the Limiting Spectral Distribution of Large-Dimensional Sample Covariance Matrices. *Annals of Probability*, **26**, 316-345. <https://doi.org/10.1214/aop/1022855421>
- [3] Bai, Z.D. and Silverstein, J.W. (2004) CLT for Linear Spectral Statistics of Large-Dimensional Sample Covariance Matrices. *Annals of Probability*, **32**, 553-605. <https://doi.org/10.1214/aop/1078415845>
- [4] Bai, Z.D., Silverstein, J.W. and Yin, Y.Q. (1998) A Note on the Largest Eigenvalue of a Large Dimensional Sample Covariance Matrix. *Journal of Multivariate Analysis*, **26**, 166-168. [https://doi.org/10.1016/0047-259X\(88\)90078-4](https://doi.org/10.1016/0047-259X(88)90078-4)
- [5] Bai, Z.D. and Yin, Y.Q. (1993) Limit of the Smallest Eigenvalue of a Large Dimensional Sample Covariance Matrix. *Annals of Probability*, **21**, 1275-1294. <https://doi.org/10.1214/aop/1176989118>
- [6] Fey-den Boer, A., van der Hofstad, R. and Klok, M.J. (2008) Large Deviations for Eigenvalues of Sample Covariance Matrices with Applications to Mobile Communication Systems. *Advance in Applied Probability*, **40**, 1048-1071. <https://doi.org/10.1017/S0001867800002962>
- [7] Guionnet, A. (2002) Large Deviation Asymptotics for Spherical Integrals. *Journal of Functional Analysis*, **188**, 461-515. <https://doi.org/10.1006/jfan.2001.3833>
- [8] Hiai, F. and Petz, D. (1998) Eigenvalue Density of the Wishart Matrix and Large Deviations. *Infinite Dimensional Analysis, Quantum Probability*, **1**, 633-646. <https://doi.org/10.1142/S021902579800034X>
- [9] Fey-den Boer, A., van der Hofstad, R. and Klok, M.J. (2003) Linear Interference Cancellation in CDMA Systems and Large Deviations of the Correlation Matrix Eigenvalues. *10th Symposium on Communications and Vehicular Technology in the Benelux*, Eindhoven University of Technology, The Netherlands, 13 November 2003.
- [10] Klok, M.J. (2001) Performance Analysis of Advanced Third Generation Receivers. Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands.
- [11] 严加安, 彭实戈, 方诗赞, 等. 随机分析选讲[M]. 北京: 科学出版社, 1997.