多天线高斯信道的容量

zcl.space

目录

多天线高斯信道的容量

1   系统模型

$$\mathbf{y} = H\mathbf{x} + \mathbf{n} \tag{1.1}$$

where H is a $r \times t$ matrix and $\mathbf{n}$ is zero-mean complex Gaussian noise with independent, equal variance real and imaginary parts. We assume $\mathbf{nn^T} = I_r$, that is the noises corrupting the different receivers are independent. The transmitter is constrained in its total power to P ,

$$\{\mathbf{x}^T\mathbf{x}\} \leq P \tag{1.2}$$

因为 $\mathbf{x}^T\mathbf{x} = tr(\mathbf{xx}^T)$, and expectation and trace commute,

$$tr(\mathcal{E}\{\mathbf{xx}^T\}) \leq P \tag{1.3}$$

可以看出有功率限制和噪声分布限制。信道矩阵限制 H ，分三种情况:

1. H is deterministic.

2. H is a random matrix (for which we shall use the notation $\boldsymbol{H}$) , chosen according to a probability distribution, and each use of the channel corresponds to an independent realization of $\boldsymbol{H}$.

3. H is a random matrix, but is fixed once it is chosen.

    if $Q \in \mathbb{C}^{n \times n}$ is non-negative devinite then so is $\hat{Q} \in \mathbb{R}^{2n \times 2n}$.

The probability density(with respect to the standard Lebesgue measure on $\mathbb{C}^n$) of a circularly symmetric complex Gaussian with mean μ and covariance Q is given by

$$\begin{aligned} \gamma_{\mu,Q} &= \det(\pi\hat{Q})^{-1/2} \exp(-(\hat{x} - \hat{\mu})^T \hat{Q}^{-1}(\hat{x} - \hat{\mu})) \\ &= \det(\pi Q)^{-1/2} \exp(-(x - \mu)^T Q^{-1}(x - \mu)) \end{aligned} \tag{1.4}$$

The differential entropy of a complex Gaussian $\mathbf{x}$ with covariance Q is given by

$$\begin{aligned} \mathcal{H}(\gamma_Q) &= \mathcal{E}_{\gamma_Q}[-\log \gamma_Q(\mathbf{x})] \\ &= \log\det(\pi Q) + (\log e)\mathcal{E}[\mathbf{x}^T Q^{-1}\mathbf{x}] \\ &= \log\det(\pi Q) + (\log e)tr(\mathcal{E}[\mathbf{xx}^T]Q^{-1}) \\ &= \log\det(\pi Q) + (\log e)tr(I) \\ &= \log\det(\pi e Q) \end{aligned} \tag{1.5}$$

For us, the importance of the circulary symmetric complex Gaussians is due to the following lemma: circularly symmetric complex Gaussians are entropy maximizers.

定理 *1.1   Suppose the complex random vector $\mathbf{x} \in \mathbb{C}^n$ is zero-mean and satisfied $\mathcal{E}[\mathbf{xx}^T] = Q$, i.e., $\mathcal{E}[\mathbf{x}_i\mathbf{x}_j^*] = Q_{ij}, 1 \leq i,j \leq n$ then the ectropy of $\mathbf{x}$ satisfies $\mathcal{H} \leq \log\det(\pi e Q)$ with equality if and only if $\mathbf{x}$ is a circulary symmetric complex Gaussian with*

$$\mathcal{E}[\mathbf{xx}^T] = Q \tag{1.6}$$

证   *Let p be any density function satisfying $\int_{\mathbb{C}^n} p(x)x_i x_j^* dx = Q_{ij}, 1 \leq j, j \leq n$ Let*

$$\gamma_Q(x) = \det(\pi Q)^{-1} \exp(-x^T Q^{-1}x). \tag{1.7}$$

Observe that $\int_{\mathbb{C}^n} \gamma_Q(x) x_i x_j^* dx = Q_{ij}$, and that $\log \gamma_Q(x)$ is a linear combination of the terms $x_i x_j^*$. Thus $\mathcal{E}_{\gamma Q}[\log \gamma_Q(\mathbf{x})] = \mathcal{E}_p[\log \gamma_Q(\mathbf{x})]$. Then,

$$
\begin{aligned}
\mathcal{H}_p - \mathcal{H}_{\gamma Q} &= -\int_{\mathbb{C}^n} p(x) \log p(x) dx + \int_{\mathbb{C}^n} \gamma_Q(x) \log \gamma_Q(x) dx \\
&= -\int_{\mathbb{C}^n} p(x) \log p(x) dx + \int_{\mathbb{C}^n} p(x) \log \gamma_Q(x) dx \\
&= \int_{\mathbb{C}^n} p(x) \log \frac{\gamma_Q(x)}{p(x)} dx \\
&\le 0,
\end{aligned}
\tag{1.8}
$$

with equality only if $p = \gamma_Q$. Thus $\mathcal{H}_p \le \mathcal{H}_{\gamma Q}$ □

定理 *1.2   if $\mathbf{x} \in \mathbb{C}^n$ is a circularly symmetric complex Gaussian then so is $\mathbf{y} = A\mathbf{x}$, for any $A \in \mathbb{C}^{m \times n}$.*

证   *We may assume $\mathbf{x}$ is zero-mean. Let $Q = \mathcal{E}[\mathbf{x}\mathbf{x}^T]$. Then y is zero-mean, $\hat{\mathbf{y}} = \hat{A}\hat{\mathbf{x}}$, and*

$$
\mathcal{E}[\hat{\mathbf{y}}\hat{\mathbf{y}}^T] = \hat{A}\mathcal{E}[\hat{\mathbf{x}}\hat{\mathbf{x}}^T]\hat{A}^T = \frac{1}{2}\hat{A}\hat{Q}\hat{A}^T = \frac{1}{2}\hat{K}.
\tag{1.9}
$$

where $K = \hat{A}\hat{Q}\hat{A}^T$. □

定理 *1.3   if $\mathbf{x}$ and $\mathbf{y}$ are independent circularly symmetric complex Gaussians, then $\mathbf{z} = \mathbf{x} + \mathbf{y}$ is a circulary symmetric complex Gaussian.*

证   *Let $A = \mathcal{E}[\mathbf{x}\mathbf{x}^T]$ and $B = \mathcal{E}[\mathbf{y}\mathbf{y}^T]$. Then $\mathcal{E}[\hat{\mathbf{z}}\hat{\mathbf{z}}^T] = \frac{1}{2}\hat{C}$ with $C = A + B$* □

2   具有固定传输函数的确定高斯信道

We will first derive an expression for the capacity $C(H, P)$ of this channel, where $H \in \mathbb{C}^{r \times t}$. By the SVD theorem $\mathbf{y} = H\mathbf{x} + \mathbf{n}$ can be written as $\mathbf{y} = UDV^T\mathbf{x} + \mathbf{n}$, where $D \in \mathbb{R}^{r \times t}$ .Let $\tilde{\mathbf{y}} = U^T\mathbf{y}, \tilde{\mathbf{x}} = V^T\mathbf{x}, \tilde{\mathbf{n}} = U^T\mathbf{n}$ .Then, $\tilde{\mathbf{y}} = D\tilde{\mathbf{x}} + \tilde{\mathbf{n}}$. Since H is of rank at most $\min\{r, t\}$, at most $\min\{r, t\}$ of the singular values of it are non-zero. Denoting these by $\lambda_i^{1/2}, i = 1, 2, \ldots, \min\{r, t\}$, we can write the matrix form of $\tilde{\mathbf{y}} = D\tilde{\mathbf{x}} + \tilde{\mathbf{n}}$ component-wise, to get $\tilde{y}_i = \lambda_i^{1/2}\tilde{x}_i + \tilde{n}_i, \quad 1 \le i \le \min\{r, t\}$, and the rest of the components of $\tilde{y}$ (if any)are equal to the corresponding components of $\tilde{n}$. We thus see that $\tilde{y}_i$ for $i > \min\{r, t\}$ is independent of the transmitted signal and that $\tilde{x}_i$ for $i > \min\{r, t\}$ don't play any role. To maximize the mutual information, we need to choose $\{\tilde{x}_i : 1 \le i \le \min\{r, t\}$ to be independent, with each $\tilde{x}_i$ having independent Gaussian, zero-mean real and imaginary parts. The variances need to be chosen via"water-filling" as

$$
\mathcal{E}[\mathfrak{Re}(\tilde{x}_i)^2] = \mathcal{E}[\mathfrak{Im}(\tilde{x}_i)^2] = \frac{1}{2}(\mu - \lambda_i^{-1})^+
\tag{2.1}
$$

where μ is chosen to meet the power constraint. Here a^+ denotes $\max\{0, a\}$. The power P and the maximal mutual information can thus be parameterized as

$$
P(\mu) = \sum_i (\mu - \lambda_i^{-1})^+, \qquad C(\mu) = \sum_i (\ln(\mu\lambda_i))^+.
\tag{2.2}
$$

2.0.1   Alternative Derivation of the capacity

The mutual information $\mathcal{I}(\mathbf{x}; \mathbf{y})$ can be written as

$$
\mathcal{I}(\mathbf{x}; \mathbf{y}) = \mathcal{H}(\mathbf{y}) - \mathcal{H}(\mathbf{y}|\mathbf{x}) = \mathcal{H}(\mathbf{y}) - \mathcal{H}(\mathbf{n})
\tag{2.3}
$$

and thus maximizing $\mathcal{I}(\mathbf{x}; \mathbf{y})$ is equivalent to maximizing $\mathcal{H}(\mathbf{y})$ Note that if $\mathbf{x}$ satisfies $\mathcal{E}\mathbf{x}\mathbf{x}^T \le P$, so does $x - \mathcal{E}[x]$, so we can restrict our attention to zero-mean $\mathbf{x}$. Furthermore, if $\mathbf{x}$ is zero-mean with covariance $\mathcal{E}[\mathbf{x}\mathbf{x}^T] = Q$, then $\mathbf{y}$ is zero-mean with covariance $\mathcal{E}[\mathbf{y}\mathbf{y}^T] = HQH^T + I_r$, when the $\mathbf{x}$ is circular symmetric complex Gaussian, the mutual information is given by:

$$
\mathcal{I}(\mathbf{x}; \mathbf{y}) = \log \det(I_r + HQH^T) = \log \det(I_r + QH^TH)
\tag{2.4}
$$

where the second equality follows from the determinant identity $\det(I + AB) = \det(I + BA)$, and it only remains to choose Q to maximize this quantity subject to the constraints $tr(Q) \leq P$ and that Q is non-negative definite. Let $\log \det(I + HQH^T) = \Psi(Q, H)$

2.1 Error Exponents

仅仅知道信道容量是不够的，通常还需要知道如何才能逼近这个容量，或者说，逼近这个容量的难度如何。于是，提出错误指数函数的概念给出差错概率一个上限。随机编码的上限是：

$$\mathcal{P}(error) \leq \exp(-nE_r(R)) \tag{2.5}$$

where the 随机编码的指数上限为：

$$E_r(R) = \max_{0 \leq \rho \leq 1} E_0 \rho - \rho R \tag{2.6}$$

where, in turn, $E_0 \rho$ is given by supremum over all input distributions q_x satisfying the energy constraint of

$$E_0(\rho, q_x) = -\log \int [\int q_x(x) p(y|x)^{1/(1+p)} dx]^{1+\rho} dy \tag{2.7}$$

In our case $p(y|x) = \det(\pi I_r)^{-1} \exp(-(y-x)^T(y-x))$. If we choose q_x as the Gaussian distribution γ_Q we get (after some algebra).

$$E_0(\rho, Q) = \rho \log \det(I_r + (1+\rho)^{-1} HQH^T) = \rho \Psi((1+\rho)^{-1}Q, H) \tag{2.8}$$

2.2 Conclusion

The use of multiple antennas will greatly increase the achievable rates on fading channels if the channel parameters can be estimated at the reciver and if the path gains between different antenna pairs behave independently. The second of these requirements can be met with relative ease and is somewhat technical in nature. The first requirement is a rather tall order, and can be justified in certain communication scenarios and not in others. Since the original writing of this monograph in late 1994 and early 1995, there has been some work in which the assumption of the availability of channel state information is replaced with the assumption of a slowly varying channel.