ECE 830 Fall 2010 Statistical Signal Processing

instructor: R. Nowak , **scribe:** Yang Song

Lecture 11: Signal Detection in Unknown Noise

Recall last lecture

$$H_0 : X \sim \mathcal{N}(0, \sigma^2 I)$$
$$H_1 : X \sim \mathcal{N}(H\theta, \sigma^2 I)$$

with σ^2 known, $H_{n\times k}$ known, and $\theta_{k\times 1}$ unknown. GLRT:

$$2\log\hat{\Lambda}(x)\frac{x^T P_H x}{2\sigma^2} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$$

Under H_0,

$$\frac{x^T P_H x}{2\sigma^2} \sim \chi_k^2$$

Example 1 *Generalize the above question:*

$$H_0 : X \sim \mathcal{N}(0, \Sigma)$$
$$H_1 : X \sim \mathcal{N}(H\theta, \Sigma)$$

with Σ known, $H_{n\times k}$ known, and $\theta_{k\times 1}$ unknown.

1 Wilk's Theorem

Theorem 1 *Suppose H_0 and H_1 composite with*

$$\underset{l\ dofs}{H_0\ models} \subset \underset{k>l\ dofs}{H_1\ models}, \quad dof: degree\ of\ freedom$$

Then under mild regularity assumptions, if

$$X_1, \dots, X_n \overset{i.i.d.}{\sim} \chi^2_{k\text{-}l}$$

Then, under H_0

$$\underbrace{2\log\hat{\Lambda}_n(x)}_{\log\ GLRT} \overset{n\to\infty}{\sim} \chi^2_{k-l}$$

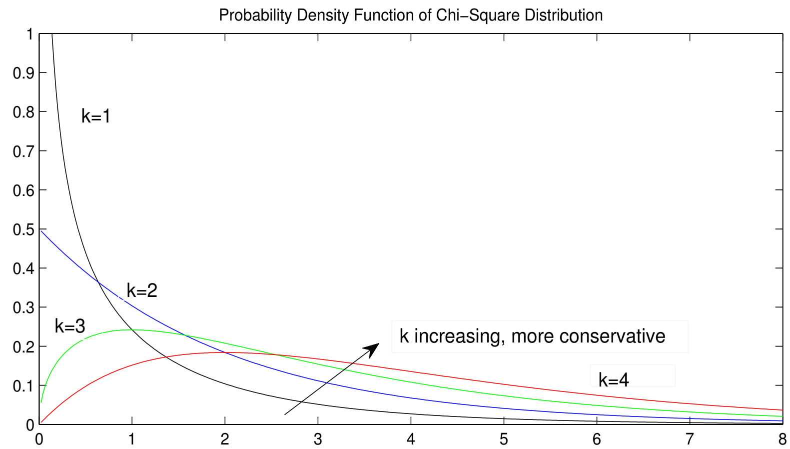

Figure 1: Probability density function of χ_k^2

2 Unknown Noise Level

Now let's look at case where noise level is unknown. Suppose

$$\begin{aligned} H_0 &: X \sim \mathcal{N}(0, \sigma^2 I) \\ H_1 &: X \sim \mathcal{N}(s, \sigma^2 I) \end{aligned}$$

where $\sigma^2 > 0$ is unknown and s is $n \times 1$ and known.

log Likelihood Ratio:

$$\log \Lambda(x) = -\frac{1}{2\sigma^2}(x-s)^T(x-s) + \frac{1}{2\sigma^2}x^T x$$

So our test is equivalent to

$$\frac{1}{\sigma^2} s^T x \underset{H_0}{\overset{H_1}{\gtrless}} \gamma'$$

or

$$t(x) := s^T x \underset{H_0}{\overset{H_1}{\gtrless}} \gamma\,, \text{ since } \sigma^2 > 0$$

Then what is the distribution of $t(x)$?

$$\begin{aligned} H_0 &: t(x) \sim \ N(0, \sigma^2 s^T s) \\ H_1 &: t(x) \sim \ N(s^T s, \sigma^2 s^T s) \end{aligned}$$

Both distributions depend on unknown σ^2!

Let's look at the GLRT. The MLE for σ^2 is

$$\hat{\sigma}_i^2 = \arg\max_{\sigma^2} \mathbb{P}(x|H_i),\ i = 0, 1$$

For H_0 we have

$$\begin{aligned}\hat{\sigma}_0^2 &= \arg\max_{\sigma^2} \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\right) x^T x \\ &= \arg\max -\frac{n}{2}\left(\log\sigma^2 + \log 2\pi\right) - \frac{1}{2\sigma^2} x^T x \\ &= \arg\min \frac{n}{2}\log\sigma^2 + \frac{1}{\sigma^2} x^T x\end{aligned}$$

Take derivative with respect to σ^2

$$\Rightarrow \frac{\partial}{\partial\sigma^2}\mathbb{P}(x|H_0) = \frac{n}{2}\frac{1}{\sigma^2} - \frac{1}{2\sigma^4}x^T x = 0$$

$$\Rightarrow \boxed{\hat{\sigma}_0^2 = \frac{1}{n}x^T x}$$

Similarly,

$$\boxed{\hat{\sigma}_1^2 = \frac{1}{n}(x-s)^T(x-s)}$$

So the GLRT is

$$\hat{\Lambda}(x) = \frac{\frac{1}{(2\pi\hat{\sigma}_1^2)^{n/2}} \exp\left(-\frac{1}{2\hat{\sigma}_1^2}(x-s)^T(x-s)\right)}{\frac{1}{(2\pi\hat{\sigma}_0^2)^{n/2}} \exp\left(-\frac{1}{2\hat{\sigma}_0^2}x^T x\right)} = \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2}\right)^{n/2}$$

the log GLRT is

$$\log\hat{\Lambda}(x) = \frac{n}{2}\log\left(\frac{x^T x}{(x-s)^T(x-s)}\right)$$

What is the distribution of X under H_0? $\quad X \sim \mathcal{N}(0, \sigma^2 I)$

$$\Rightarrow \underset{\substack{\uparrow \\ \textit{still a function of } \sigma^2!}}{\log\hat{\Lambda}(x)} = \frac{n}{2}\log\left(\frac{\sigma^2 w^T w}{(\sigma w - s)^T(\sigma w - s)}\right), \quad \textit{where } w \sim \mathcal{N}(0, \sigma^2 I)$$

And, Wilks' Theorem doesn't apply since both H_0 and H_1 models have one degree of freedom. We cannot set γ to control either error.

3 Unknown Signal and Noise Amplitudes

Let's look at a slightly different problem. The problem in the previous case is that the unkown noise amplitude affected the variance of both distributions, and the MLE of the noise variance differed in the two hypotheses. Let us now suppose:

$$\begin{aligned}H_0 &: X \sim \mathcal{N}(0, \sigma^2 I) \\ H_1 &: X \sim \mathcal{N}(\theta s, \sigma^2 I)\end{aligned}$$

with σ^2 unknown, θ unknown, and $s_{n\times 1}$ unknown.
Under H_0

$$X = \sigma w, \quad w \sim \mathcal{N}(0, I)$$

Under H_1

$$\begin{aligned} X = & \ \theta s + \sigma w \\ = & \ \sigma(\theta' s + w), \quad where \quad \theta' = \frac{\theta}{\sigma} \end{aligned}$$

The advantage here is that σ can be viewed as a scaling factor for the observation in both cases.

Let's consider the GLRT for this problem. Under H_0

$$\hat{\sigma}_0^2 = \frac{1}{n} x^T x, \quad asbefore.$$

Under H_1, we must find the MLE of θ and σ^2.

$$\mathbb{P}(x|H,\ \theta,\ \sigma^2) = \frac{1}{(s\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2sigma^2}(x-\theta s)^T(x-\theta s)\right)$$

Taking the log we have:

$$-\frac{n}{2}(\log\sigma^2 + \log 2\pi) - \frac{1}{2\sigma^2}(x^T x - 1\theta s^T x + \theta^2 s^T s)$$

Differentiating with respect to σ^2:

$$\begin{aligned} s^T x = & \ \theta s^T s \\ \Rightarrow \boxed{\hat{\theta} = \frac{s^T x}{s^T s}} \end{aligned}$$

Differentiating with respect to σ^2

$$\begin{aligned} -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma}(x-\hat{\theta}s)^T(x-\hat{\theta}s) = & \ 0 \\ \Rightarrow \boxed{\hat{\sigma}_1^2 = \frac{1}{n}(x-\hat{\theta}s)^T(x-\hat{\theta}s)} \end{aligned}$$

Note: Wilks' Theorem applies:

$$2\log\frac{\mathbb{P}(X|H_1,\hat{\theta},\hat{\sigma}_1^2)}{\mathbb{P}(X|H_0,\hat{\sigma}_0^2)} \sim \chi_1^2 \quad for\ large\ n$$

Let's look at the GLRT more closely.

$$\begin{aligned} \hat{\Lambda}(x) = & \ \frac{\mathbb{P}(x|H_1,\hat{\theta},\hat{\sigma^2})}{\mathbb{P}(x|H_0,\hat{\sigma_0^2})} \\ = & \ \left(\frac{\hat{\sigma}_0^2}{\sigma_1^2}\right)^{n/2} \exp\left(-\frac{1}{2\hat{\sigma}_1^2}(x-\hat{\theta}s)^T(x-\hat{\theta}s) + \frac{1}{2\hat{\sigma}_0^2 x^T x}\right) \\ = & \ \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2}\right)^{n/2} \exp\left(-\frac{n}{2}+\frac{n}{2}\right) \\ = & \ \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}_0^2}\right) \end{aligned}$$

So the GLRT has the simple form

$$\frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2} \;=\; \frac{x^T x}{(x-\hat{\theta}s)^T(x-\hat{\theta}s)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$$

or equivalently

$$\frac{x^T x}{(x-\frac{ss^Tx}{s^Ts})(x-\frac{ss^Tx}{s^Ts})} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$$

Under H_0 we have $X \;=\; \sigma w$, $w \sim \mathcal{N}(0,\; I)$, and so the test statistic is

$$\frac{\sigma^2 w^T w}{\sigma^2 (w-P_s w)^T(w-P_s w)} \;\rightarrow\; \text{invariant to } \sigma^2$$

Definition 1 *γ can be chosen to insure a special P_{FA} for every value of σ^2. A test like this is said to have a constant false alarm rate and is called CFAR detector.*

To set γ we need to determine the distribution of $\frac{w^T w}{(w-P_s w)^T(w-P_s w)}$, $w \sim \mathcal{N}(0, I)$. Consider the test statistic

$$\begin{aligned} t(x) &= \frac{x^T x}{(x-P_s x)^T(x-P_s x)} \\ &= \frac{x^T x}{x(I-P_s)x} = \frac{x^T(I-P_s)x + x^T P_s x}{x^T(I-P_s)x} \\ &= 1 + \frac{x^t P_s x}{x^T(I-P_s)x} \end{aligned}$$

So equivalently, we can write the GLRT as:

$$\frac{x^T P_s x}{x^T(I-P_s)x} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$$

Let U be $n \times (n-1)$ matrix whose orthonormal columns span subspace orthogonal to s.

$$U \;=\; [u_1, \ldots, u_{n-1}]$$

Then $u_1, \ldots, u_{n-1}, \frac{s}{\|s\|}$ are orthonormal basis for $\mathbb{R}^n$

$$\begin{aligned} x^T P_s x \;&=\; \frac{|s^T x|^2}{\|s\|^2} \\ x^T P_s s \;&=\; \sum_{i=1}^{n-1} |u_i^T x|^2 \end{aligned}$$

Under H_0

$$\begin{aligned} \frac{|s^T w|^2}{\|s\|^2} &\sim\; \chi_1^2 \\ \sum_{i=1}^{n-1} (u_i^T w)^2 &\sim\; \chi_{n-1}^2 \end{aligned}$$

Moreover, because $u_1, \ldots, u_{n-1}$ are orthogonal to s, the $s^T x$ and $u_i^T x,\; i = 1, \ldots, n-1$ are uncorrelated and thus independent.

$$\Rightarrow \frac{(s^T w)^2}{\|s\|^2} \text{ are independent!}$$

The ratio of independent χ^2 random variables with degree of freedom k and l respectively has been well studied and has a name: F-distributed with $(k,\ l)$ degree of freedom

$$\frac{\chi_k^2/k}{\chi_l^2/l} \ \sim\ F_{k,\ l}$$

In our case, under H_0

$$\frac{x^T P_s x}{x^T(I - P_s)x/(n-1)} \ \sim\ F_{1,\ n-1}$$

and thus we can use the tail of the F-distribution to set a threshold for a desired P_{FA}.

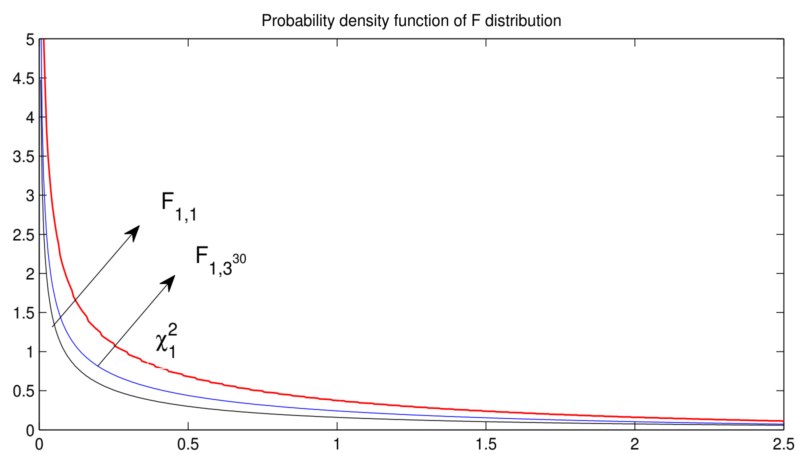

Figure 2: Probability density function of F distribution and χ_1^2 distribution

Note:
If $X \sim\ F(\nu_1, \nu_2)$, then $Y\ =\ \lim_{\nu_2\to\infty} \nu_1 X$ has the chi-square distribution $\chi_{\nu_1}^2$.
In our case, $\lim_{n\to\infty} F_{1,n-1} \sim\ \chi_1^2$, which is what Wilks' Theorem told us.