

به نام هستی بخش



دانشکده برق و کامپیوتر

انتخاب مرتبه مدل بر اساس روش های تئوری اطلاعات/شکل دهی پرتو

سمینار درس تخمین طیف

حسین واعظی

استاد: آقای دکتر نقش

تیر ۹۴

فهرست مطالب

تابع اختلاف KL

روش های انتخاب مرتبه مدل

خلاصه روش ها

تخمین فرکانس سینوسی

شبیه سازی تخمین فرکانس

Capon Beamforming

تابع اختلاف KL

$$D(p_0, \hat{p}) = \int p_0(y) \ln \left[\frac{p_0(y)}{\hat{p}(y)} \right] dy$$

$p_0(y)$: تابع چگالی احتمال واقعی داده‌ها

$\hat{p}(y)$: تابع چگالی احتمال یک مدل نوعی از داده‌ها

$$\begin{aligned} D(p_0, \hat{p}) &= E_0 \left\{ \ln \left[\frac{p_0(y)}{\hat{p}(y)} \right] \right\} \\ &= E_0 \{ \ln p_0(y) \} - E_0 \{ \ln \hat{p}(y) \} \end{aligned}$$

خواص اختلاف KL

$$\left\{ \begin{array}{l} D(p_0, \hat{p}) \geq 0 \\ D(p_0, \hat{p}) = 0 \text{ if and only if } p_0(y) = \hat{p}(y) \end{array} \right.$$

انتخاب مرتبه مدل

$$I(p_0, \hat{p}) \triangleq E_0\{\ln \hat{p}(y)\}$$

$$\min D(p_0, \hat{p}) \equiv \max I(p_0, \hat{p})$$

با فرض $\hat{p}(y) = p(y, \hat{\theta})$

$$I(p_0, p(y, \hat{\theta})) = E_0\{\ln p(y, \hat{\theta})\}$$

روش اول:
به دلیل نامعلوم بودن تابع چگالی احتمال داده ها از تخمین بدون بایاس I استفاده می کنیم.

$$\hat{I} = \ln p(y, \hat{\theta})$$

انتخاب مرتبه مدل

روش دوم:

$$\ln p_n(y, \theta^n)$$

$$\approx \ln p_n(y, \hat{\theta}^n) + (\theta^n - \hat{\theta}^n)^T \left[\frac{\partial \ln p_n(y, \theta^n)}{\partial \theta^n} \right]_{\theta^n = \hat{\theta}^n}$$

$$+ \frac{1}{2} (\theta^n - \hat{\theta}^n)^T \left[\frac{\partial^2 \ln p_n(y, \theta^n)}{\partial \theta^n \partial \theta^{nT}} \right]_{\theta^n = \hat{\theta}^n} \times (\theta^n - \hat{\theta}^n)$$

$\triangleq \ln \hat{p}_n(y)$

که ترم دوم برابر صفر می باشد.

$$\ln \hat{p}_n(y) = \ln p_n(y, \hat{\theta}^n) - \frac{1}{2} (\theta^n - \hat{\theta}^n)^T J (\theta^n - \hat{\theta}^n)$$

$$E \left\{ (\theta^n - \hat{\theta}^n)^T J (\theta^n - \hat{\theta}^n) \right\} = \text{tr} \left[J E \left\{ (\theta^n - \hat{\theta}^n)^T (\theta^n - \hat{\theta}^n) \right\} \right] = \text{tr} [I_n] = n$$

انتخاب مرتبه مدل

$$I = E \left\{ \ln p_n(y, \hat{\theta}^n) - \frac{n}{2} \right\}$$

بنابراین داریم:

تخمین بدون بایاس I برابر است با:

$$\hat{I} = \ln p_n(y, \hat{\theta}^n) - \frac{n}{2}$$

قاعده no-name:

$$\min NN(n) = \min -2 \ln p_n(y, \hat{\theta}^n) + n$$

باید تاثیر عبارت جریمه را زیاد کرد.



مشکل: **overfitting**

انتخاب مرتبه مدل

روش AIC:

مشکل ۱: در دسترس نبودن مدل درست‌نمایی $\hat{p}(y)$

مشکل ۲: محاسبه امید ریاضی $E_0\{\ln \hat{p}(y)\}$ غیر ممکن است. (تابع چگالی نامعلوم داده)

راه حل: استفاده از تخمینی برای \hat{I}

فرض: X برداری جعلی ازداده که تابع چگالی آن با تابع چگالی Y یکسان باشد.

$$\ln \hat{p}(y) = E_x\{\ln p(y, \hat{\theta}_x)\}$$

$\hat{\theta}_x$ تخمین ML بردار پارامتر مدل

انتخاب مرتبه مدل

روش AIC:

$$I = E_y \{ E_x \{ \ln p(y, \hat{\theta}_x) \} \}$$

$$I \approx E_y \{ \ln p_n(y, \hat{\theta}^n) - n \}$$

$$AIC = -2 \ln p_n(y, \hat{\theta}^n) + 2n$$

the probability of underfitting $\rightarrow 0$

the probability of overfitting $\rightarrow \text{constant} > 0$

مثال: برای مدل سیگنال سینوسی با n_c مولفه

$$AIC = 2N_s \ln \hat{\sigma}_{n_c}^2 + 2(3n_c + 1)$$

$$\hat{\sigma}_{n_c}^2 = \frac{1}{N_s} \sum_{t=1}^{N_s} \left| y_c(t) - \sum_{k=1}^{n_c} \hat{\alpha}_k e^{i(\hat{\omega}_k t + \hat{\varphi}_k)} \right|^2$$

انتخاب مرتبه مدل

روش GIC:

$$GIC = -2 \ln p_n(y, \hat{\theta}^n) + \nu n$$

با افزایش طول داده x خطر **overfitting** کاهش می یابد. بنابراین:

$$N \triangleq \text{length}(y) = \rho \cdot \text{length}(x), \quad \rho \geq 1$$

$$GIC = -2 \ln p_n(y, \hat{\theta}_y) + (1 + \rho)n$$

مشکل: انتخاب مقدار ρ

انتخاب مرتبه مدل

روش BIC:

اگر $p(\theta)$ در اطراف $\hat{\theta}$ هموار باشد.

$$\hat{p}(y) = E_{\theta}\{p(y, \theta)\} = \int p(y, \theta)p(\theta)d\theta$$

$$p(y, \theta) \approx p(y, \hat{\theta})e^{-\frac{1}{2}(\hat{\theta}-\theta)^T \hat{J}(\hat{\theta}-\theta)}$$

$$\begin{aligned}\hat{p}(y) &\approx p(y, \hat{\theta})p(\hat{\theta}) \int e^{-\frac{1}{2}(\hat{\theta}-\theta)^T \hat{J}(\hat{\theta}-\theta)} d\theta \\ &= \frac{p(y, \hat{\theta})p(\hat{\theta})(2\pi)^{n/2}}{|\hat{J}|^{1/2}}\end{aligned}$$

انتخاب مرتبه مدل

روش BIC:

$$\hat{I} = \ln p(y, \hat{\theta}) + \ln p(\hat{\theta}) + \frac{n}{2} \ln 2\pi - \frac{1}{2} \ln |\hat{J}|$$

$$\hat{I} \approx \ln p_n(y, \hat{\theta}^n) - \frac{n}{2} \ln N$$

$$BIC = -2 \ln p_n(y, \hat{\theta}^n) + n \ln N$$

انتخاب مرتبه مدل

خلاصه

$$\text{Cri} = -2 \ln p_{n_c}(y, \hat{\theta}^n) + \eta(n, N)n$$

AIC	$\eta(n, N) = 2$
AIC _c	$\eta(n, N) = 2 \frac{N}{N - n - 1}$
GIC	$\eta(n, N) = \nu = \rho + 1$
BIC	$\eta(n, N) = \ln N$

معیارهای دیگر انتخاب مرتبه

Estimator	Formula
Schwarz-Bayes Criterion (Bayesian Information Criterion)	$SBC(p) = \ln \tilde{\Sigma}(p) + \frac{\ln(\hat{T})}{\hat{T}} pM^2$
Akaike Information Criterion	$AIC(p) = \ln \tilde{\Sigma}(p) + \frac{2}{\hat{T}} pM^2$
Akaike's Final Prediction Error	$\ln(FPE(p)) = \ln \tilde{\Sigma}(p) + M \ln \left(\frac{\hat{T} + Mp + 1}{\hat{T} - Mp - 1} \right)$
Hannan-Quinn Criterion	$HQ(p) = \ln \tilde{\Sigma}(p) + \frac{2 \ln(\ln(\hat{T}))}{\hat{T}} pM^2$

شبه سازی روش Capon برای تخمین فرکانس سینوسی

$$\begin{cases} \min & h^H R h \\ \text{s.t:} & h^H a(\omega) = 1 \end{cases}$$

$$h = \frac{R^{-1} a(\omega)}{a^H(\omega) R^{-1} a(\omega)}$$

$$\varphi(\omega) = \frac{1}{\beta a^H(\omega) R^{-1} a(\omega)}$$

$$\beta = m + 1$$

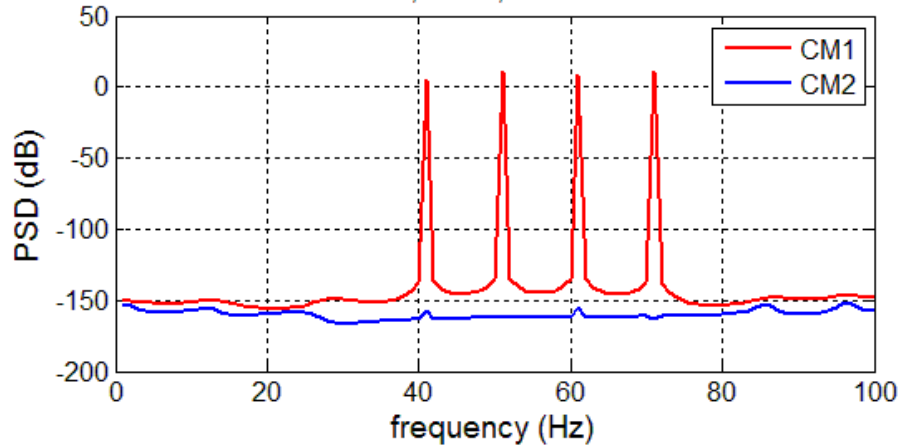
$$\beta = h^H h$$

$$\hat{\varphi}_{CM1}(\omega) = \frac{m + 1}{a^H(\omega) R^{-1} a(\omega)}$$

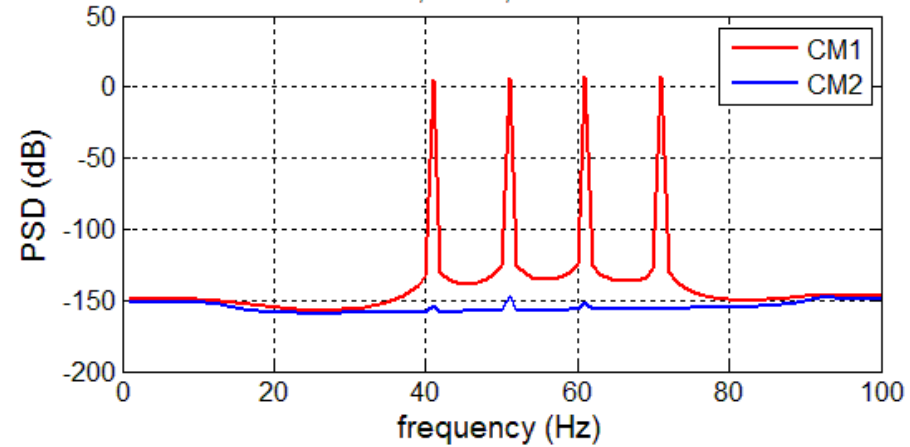
$$\hat{\varphi}_{CM2}(\omega) = \frac{a^H(\omega) \hat{R}^{-1} a(\omega)}{a^H(\omega) \hat{R}^{-2} a(\omega)}$$

شبیه سازی روش Capon برای تخمین فرکانس سینوسی

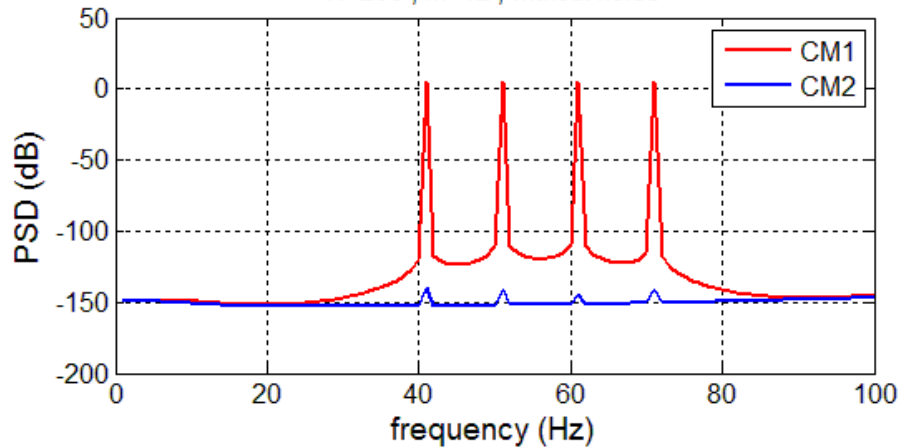
N=200 ; m=18 ; without noise



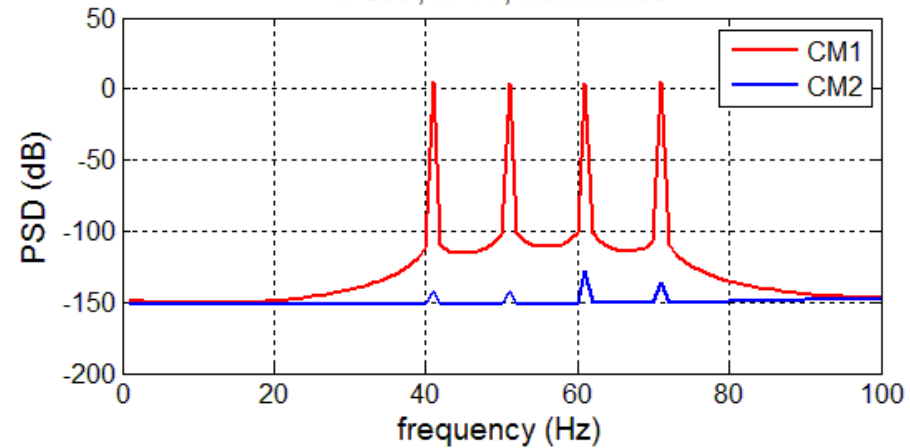
N=200 ; m=16 ; without noise



N=200 ; m=12 ; without noise

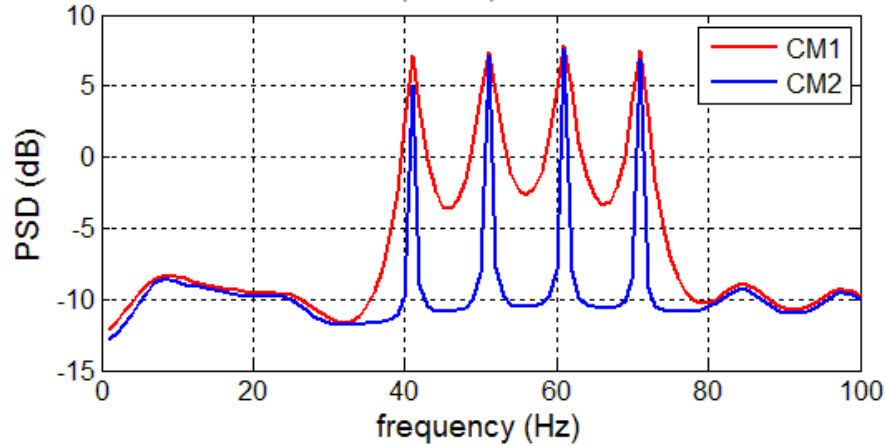


N=200 ; m=10 ; without noise

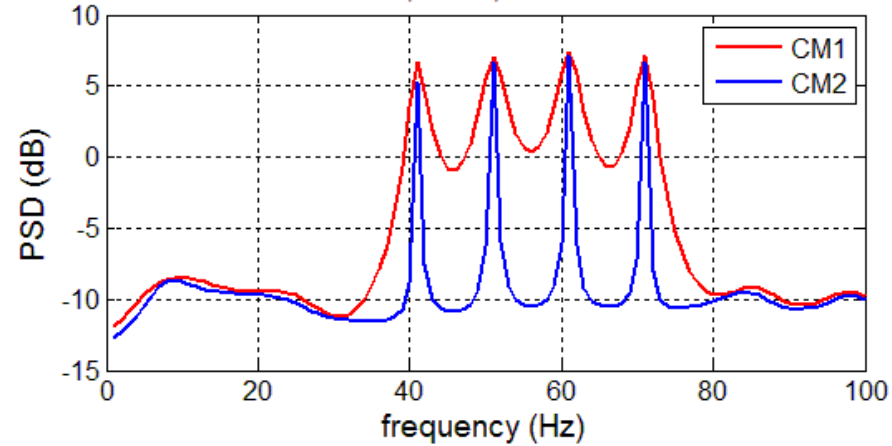


شبیه سازی روش Capon برای تخمین فرکانس سینوسی

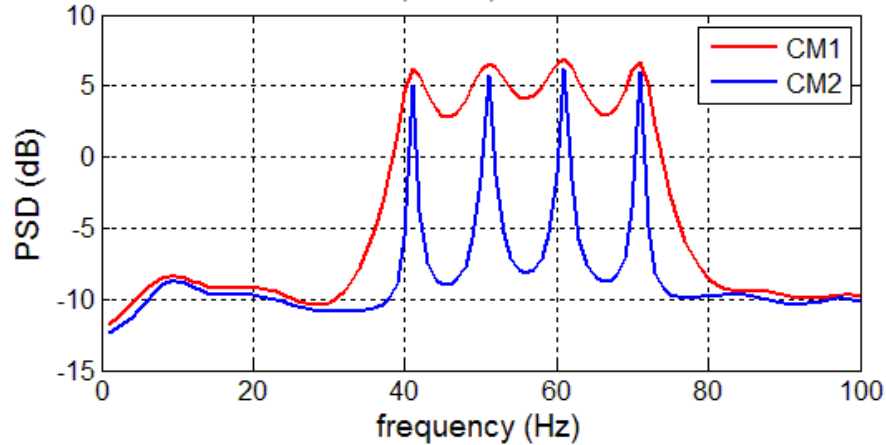
N=200 ; m=20 ; SNR=-10dB



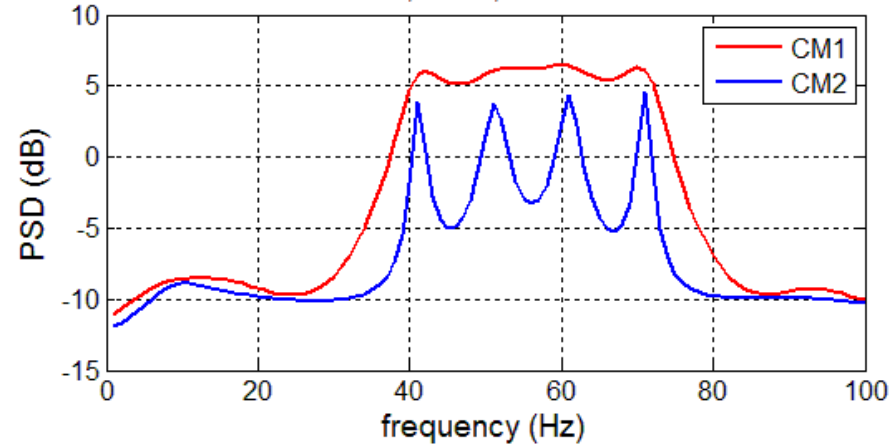
N=200 ; m=18 ; SNR=-10dB



N=200 ; m=16 ; SNR=-10dB

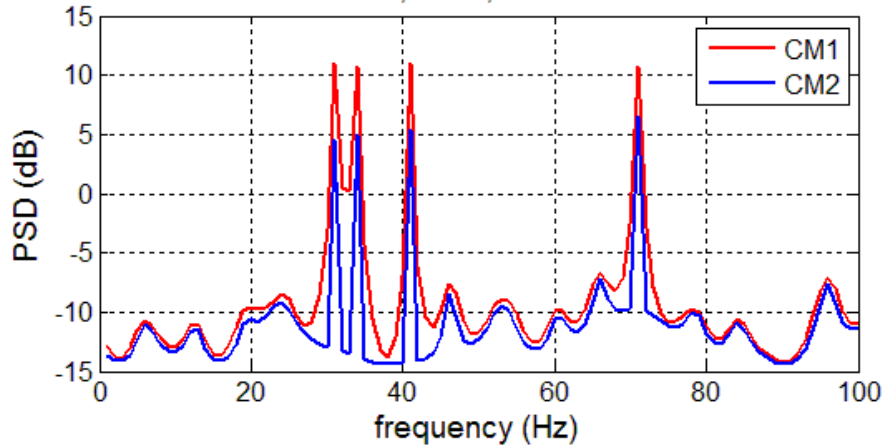


N=200 ; m=14 ; SNR=-10dB

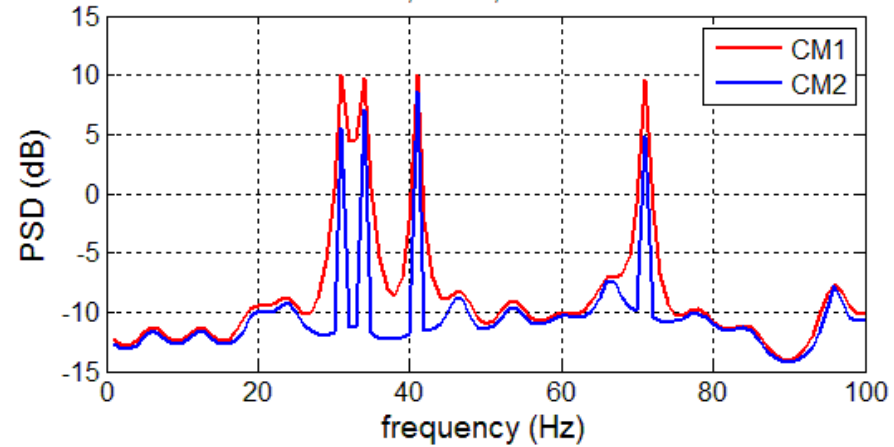


شبیه سازی روش Capon برای تخمین فرکانس سینوسی

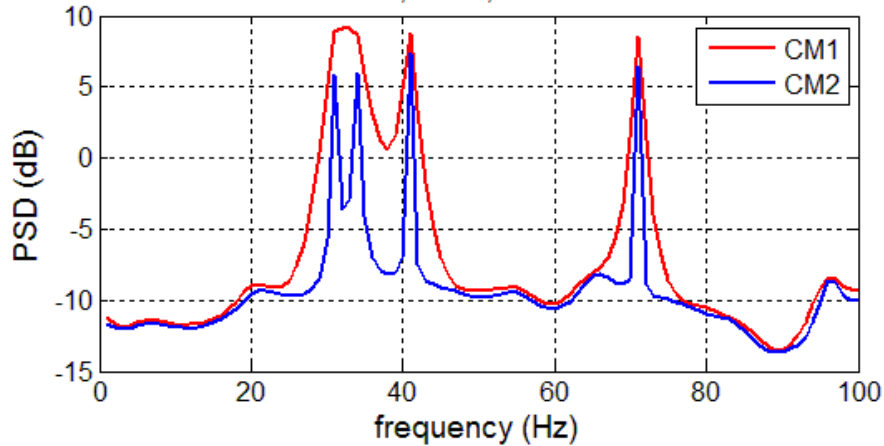
N=200 ; m=50 ; SNR=-10dB



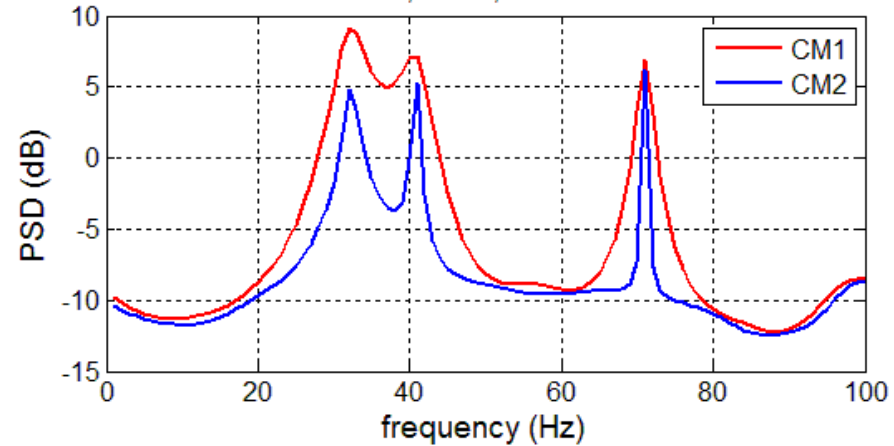
N=200 ; m=40 ; SNR=-10dB



N=200 ; m=30 ; SNR=-10dB



N=200 ; m=20 ; SNR=-10dB



با شکر از توجه شما!



