A CLT on the SINR of the diagonally loaded Capon/MVDR beamformer

Francisco Rubio[1]

joint work with **Xavier Mestre**[1] and **Walid Hachem**[2]

[1]Centre Tecnològic de Telecomunicacions de Catalunya

[2]Télécom ParisTech and Centre National de la Recherche Scientifique

Workshop on Large Random Matrices and their Applications
Télécom ParisTech, 11-13 October, 2010

Outline

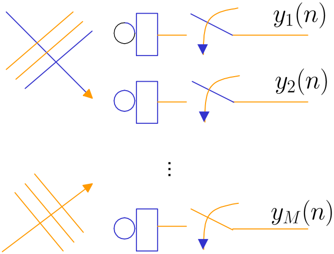
- Capon/MVDR beamforming (or spatial filtering)
- Characterization of output SINR performance
- Asymptotic deterministic equivalents
- A Central Limit Theorem
- Conclusions

Capon/MVDR beamforming
Signal model

- Consider the following set of independent observations drawn from the general Gauss-Markov linear model $\mathcal{L}\left(\mathbf{y}\left(n\right), x\left(n\right)\mathbf{s}, \mathbf{R}\right)$:

$$\mathbf{y}\left(n\right) = x\left(n\right)\mathbf{s} + \mathbf{n}\left(n\right) \in \mathbb{C}^{M}, \quad n = 1, \ldots, N$$

  where $x\left(n\right) \equiv$ **signal waveform**, $\mathbf{s} \in \mathbb{C}^{M} \equiv$ **spatial signature**, $\mathbf{n}\left(n\right) \in \mathbb{C}^{K} \equiv$ **i+n**
- Typical scenario in sensor array signal processing applications:



$$\mathbf{y}(n) = \begin{bmatrix} y_1(n) \\ y_2(n) \\ \vdots \\ y_M(n) \end{bmatrix}, n = 1 \ldots N$$

- We are interested in linearly filtering the observed samples to estimate $x\left(n\right)$

- **Optimal coefficients** of **M**inumum **V**ariance **D**istornionless **R**esponse filter:

$$\mathbf{w}_{\text{MVDR}} = \arg \min_{\mathbf{w} \in \mathbb{C}^M} \mathbf{w}^H \mathbf{R} \mathbf{w} \ \text{ subject to } \mathbf{w}^H \mathbf{s} = 1$$
$$= \frac{\mathbf{R}^{-1}\mathbf{s}}{\mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}}$$

  where $\mathbf{R}$ is the covariance matrix of interference-plus-noise random vectors

- In practice, $\mathbf{R}$ is unknown and implementations rely on the **S**ample **C**ovariance **M**atrix or any other improved estimator based on **regularization** or **shrinkage**:

$$\hat{\mathbf{R}} = \frac{1}{N} \mathbf{Y} \left(\mathbf{I}_N - \frac{1}{N} \mathbf{1}_N \mathbf{1}'_N \right) \mathbf{Y}^H + \alpha \mathbf{R}_o, \qquad \mathbf{Y} = \left[\mathbf{y}(1), \ldots, \mathbf{y}(N) \right]$$

  where $\mathbf{R}_o$ is a positive matrix and $\alpha > 0$ is the **diagonal loading** or **shrinkage intensity** parameter

- If $\alpha = 0$ then $\hat{\mathbf{R}} = \hat{\mathbf{R}}_{\text{SCM}}$ and, under Gaussianity, $\hat{\mathbf{R}}_{\text{SCM}} \overset{\mathcal{L}}{=} \frac{1}{N} \mathbf{R}^{1/2} \mathbf{X} \mathbf{T} \mathbf{X}^H \mathbf{R}^{1/2}$ where the entries of $\mathbf{X}$ are $\mathcal{CN}(0,1)$, and $\mathbf{T}$ models either **sample weighting** or **temporal correlation** across samples

- The **S**ignal-to-**I**nterference-plus-**N**oise **R**atio at the output of the MVDR filter is:

$$\mathsf{SINR}\left(\mathbf{w}\right) = \frac{\sigma_x^2 \left|\mathbf{w}^H \mathbf{s}\right|^2}{\mathbf{w}^H \mathbf{R} \mathbf{w}}$$

with $\sigma_x^2 \equiv$ **signal power**

- The optimal SINR is $\mathsf{SINR}\left(\mathbf{w}_{\mathsf{MVDR}}\right) = \mathbf{s}^H \mathbf{R}^{-1} \mathbf{s} \equiv \left\|\mathbf{u}\right\|^2$
- For the MVDR filter implementation based on diagonal loading:

$$\mathsf{SINR}\left(\hat{\mathbf{w}}_{\mathsf{MVDR}}\right) = \frac{\left(\mathbf{s}^H \left(\hat{\mathbf{R}} + \alpha \mathbf{I}_M\right)^{-1} \mathbf{s}\right)^2}{\mathbf{s}^H \left(\hat{\mathbf{R}} + \alpha \mathbf{I}_M\right)^{-1} \mathbf{R} \left(\hat{\mathbf{R}} + \alpha \mathbf{I}_M\right)^{-1} \mathbf{s}}$$

- We are interested in the properties of $\mathsf{SINR}\left(\hat{\mathbf{w}}_{\mathsf{MVDR}}\right)$

- In the case $\hat{\mathbf{R}} = \hat{\mathbf{R}}_{\mathsf{SCM}}$ ($\mathbf{T} = \mathbf{I}_N$ and $\alpha = 0$), the distribution of

$$\frac{\mathsf{SINR}\left(\hat{\mathbf{w}}_{\mathsf{MVDR}}\right)}{\mathsf{SINR}\left(\mathbf{w}_{\mathsf{MVDR}}\right)} = \frac{\left(\mathbf{s}^H \hat{\mathbf{R}}^{-1} \mathbf{s}\right)^2}{\mathbf{s}^H \hat{\mathbf{R}}^{-1} \mathbf{R} \hat{\mathbf{R}}^{-1} \mathbf{s} \mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}}$$

  is known in the array processing literature to have a density [**Reed-Mallet-Brennan,T.AES'74**]

$$f_\rho\left(\rho\right) = \frac{N!}{(M-2)!\,(N+1-M)!}\left(1-\rho\right)^{M-2}\rho^{N+1-M}$$

- In particular, $\mathsf{SINR}\left(\hat{\mathbf{w}}_{\mathsf{MVDR}}\right)/\mathsf{SINR}\left(\mathbf{w}_{\mathsf{MVDR}}\right) \sim \mathrm{Beta}\left(N+2-M, M-1\right)$ with

$$\mathrm{mean} = \frac{N+2-M}{N+1}$$

  and

$$\mathrm{variance} = \frac{\left(M-1\right)\left(N+2-M\right)}{\left(N+1\right)^2\left(N+2\right)}$$

- What about the general case with arbitrary positive $\mathbf{T}$ and α ? [**Rao-Edelman, ASAP'05**]

- First-order analysis:

$$\mathsf{SINR}\left(\hat{\mathbf{w}}_{\mathsf{MVDR}}\right) = \frac{\left(\mathbf{s}^H\left(\hat{\mathbf{R}}+\alpha\mathbf{I}_M\right)^{-1}\mathbf{s}\right)^2}{\mathbf{s}^H\left(\hat{\mathbf{R}}+\alpha\mathbf{I}_M\right)^{-1}\mathbf{R}\left(\hat{\mathbf{R}}+\alpha\mathbf{I}_M\right)^{-1}\mathbf{s}}$$

$$\asymp \frac{\left(\mathbf{s}^H\left(x_M\mathbf{R}+\alpha\mathbf{R}_o\right)^{-1}\mathbf{s}\right)^2}{\frac{1}{1-\gamma\tilde{\gamma}}\mathbf{s}^H\left(x_M\mathbf{R}+\alpha\mathbf{R}_o\right)^{-1}\mathbf{R}\left(x_M\mathbf{R}+\alpha\mathbf{R}_o\right)^{-1}\mathbf{s}} = \overline{\mathsf{SINR}\left(\hat{\mathbf{w}}_{\mathsf{MVDR}}\right)}$$

such that

$$x_M = \frac{1}{N}\operatorname{Tr}\left[\mathbf{T}\left(\mathbf{I}_N+e_M\mathbf{T}\right)^{-1}\right] \equiv \frac{1}{N}\operatorname{Tr}\left[\tilde{\mathbf{E}}\right]$$

$$e_M = \frac{1}{N}\operatorname{Tr}\left[\mathbf{R}\left(x_M\mathbf{R}+\alpha\mathbf{R}_o\right)^{-1}\right] \equiv \frac{1}{N}\operatorname{Tr}\left[\mathbf{E}\right]$$

and $\gamma = \frac{1}{N}\operatorname{Tr}\left[\mathbf{E}^2\right]$ and $\tilde{\gamma} = \frac{1}{N}\operatorname{Tr}\left[\tilde{\mathbf{E}}^2\right]$

- Asymptotics of $\mathsf{SINR}\left(\hat{\mathbf{w}}_{\mathsf{MVDR}}\right)$ involve **both the eigenvalues and also the eigenvectors** of the random matrix model

- If the entries of $\mathbf{X}$ have $8th$-order moment and $\|\mathbf{R}\|$ and $\|\mathbf{T}\|$ are bounded, as $N = N(M) \to \infty$ and $0 \leq \liminf c_M \leq \limsup c_M < +\infty$ ($c_M = M/N$), a.s., [**Rubio-Mestre, submitted SPL'10**]

$$\upsilon^H \left(\mathbf{A} + \tfrac{1}{N}\mathbf{R}^{1/2}\mathbf{X}\mathbf{T}\mathbf{X}^H\mathbf{R}^{1/2} - z\mathbf{I}_M\right)^{-1} \upsilon \asymp \upsilon^H \left(\alpha\mathbf{R}_o + x(z)\mathbf{R} - z\mathbf{I}_M\right)^{-1} \upsilon$$

for each $z \in \mathbb{C} - \mathbb{R}^+$ and an arbitrary nonrandom, unit-norm υ, where

$$x(z) = \frac{1}{N}\operatorname{Tr}\left[\mathbf{T}\left(\mathbf{I}_N + e(z)\mathbf{T}\right)^{-1}\right]$$

and $e(z)$ is the unique solution in $\mathbb{C} - \mathbb{R}^+$ to

$$e(z) = \frac{1}{N}\operatorname{Tr}\left[\mathbf{R}\left(\alpha\mathbf{R}_o + x(z)\mathbf{R} - z\mathbf{I}_M\right)^{-1}\right]$$

- Define $\mathbf{Q}_M(z) = \left(\tfrac{1}{N}\mathbf{X}\mathbf{T}\mathbf{X}^H + \alpha\mathbf{R}^{-1} - z\mathbf{I}_M\right)^{-1}$ and note that $\mathbf{Q}_M^2(z) = \frac{\partial}{\partial z}\left\{\mathbf{Q}_M(z)\right\}\big|_{z=0}$ along with

$$\mathsf{SINR}\left(\hat{\mathbf{w}}_{\mathsf{MVDR}}\right) = \frac{\left(\mathbf{u}^H\mathbf{Q}_M(0)\,\mathbf{u}\right)^2}{\mathbf{u}^H\mathbf{Q}_M^2(0)\,\mathbf{u}}$$

- We also have the following estimate not depending on the unknown $\mathbf{R}$:

$$\mathsf{SINR}\left(\hat{\mathbf{w}}_{\mathsf{MVDR}}\right) \asymp \phi_M\left(\alpha\right) \times \frac{\mathbf{s}^H\left(\hat{\mathbf{R}} + \alpha \mathbf{R}_o\right)^{-1}\hat{\mathbf{R}}\left(\hat{\mathbf{R}} + \alpha \mathbf{R}_o\right)^{-1}\mathbf{s}}{\left(\mathbf{s}^H\left(\hat{\mathbf{R}} + \alpha \mathbf{R}_o\right)^{-1}\mathbf{s}\right)^2}$$

where

$$\phi_M\left(\alpha\right) = \frac{1}{1 - \frac{1}{N}\operatorname{Tr}\left[\hat{\mathbf{R}}\left(\hat{\mathbf{R}} + \alpha \mathbf{R}_o\right)^{-1}\right]}$$

- The previous estimate can be used to find the **optimal diagonal loading factor** or shrinkage intensity parameter for arbitrary shrinkage target $\mathbf{R}_o$
- What about the fluctuations of $\mathsf{SINR}\left(\hat{\mathbf{w}}_{\mathsf{MVDR}}\right)$?

- We analyze the variance σ_M^2 of $\mathsf{SINR}\left(\hat{\mathbf{w}}_{\mathsf{MVDR}}\right)$ and prove the Central Limit Theorem

$$\sigma_M^{-1}\left(\mathsf{SINR}\left(\hat{\mathbf{w}}_{\mathsf{MVDR}}\right)-\overline{\mathsf{SINR}\left(\hat{\mathbf{w}}_{\mathsf{MVDR}}\right)}\right)\underset{M,N\to\infty}{\overset{\mathcal{L}}{\to}}\mathcal{N}\left(0,1\right)$$

by applying the **Delta method** to the random vector

$$\left[\begin{array}{c} a_M \\ b_M \end{array}\right]=\left[\begin{array}{c} \mathbf{s}^H\left(\hat{\mathbf{R}}+\alpha\mathbf{I}_M\right)^{-1}\mathbf{s} \\ \mathbf{s}^H\left(\hat{\mathbf{R}}+\alpha\mathbf{I}_M\right)^{-1}\mathbf{R}\left(\hat{\mathbf{R}}+\alpha\mathbf{I}_M\right)^{-1}\mathbf{s} \end{array}\right]$$

whose distribution is obtained by using the **Cramér-Wold device** after managing the following computations...

- Recall $\mathbf{Q}_M(0) = \mathbf{Q}_M = \left(\frac{1}{N}\mathbf{X}\mathbf{T}\mathbf{X}^H + \alpha\mathbf{R}^{-1}\right)^{-1}$ and $\|\mathbf{u}\|^2 = \mathbf{s}^H\mathbf{R}^{-1}\mathbf{s}$, and define

$$\bar{a}_M \asymp a_M = \mathbf{u}^H\mathbf{Q}_M\mathbf{u}$$
$$\bar{b}_M \asymp b_M = \mathbf{u}^H\mathbf{Q}_M^2\mathbf{u}$$

- We follow the approach by Hachem *et al.* in [**H-K-L-N-P, T.IT'2008**] and show that

$$\Psi_M(\omega) - \exp\left(-\omega^2\sigma_M^2/2\right) \underset{M,N\to\infty}{\longrightarrow} 0$$

where $\Psi_M(\omega)$ is the **characteristic function** of the random variable

$$A\sqrt{N}\left(a_M - \bar{a}_M\right) - B\sqrt{N}\left(b_M - \bar{b}_M\right)$$

- To identify the variance, we proceed as

$$\frac{\partial}{\partial\omega}\Psi_M(\omega) = \mathrm{i}\,A\sqrt{N}\mathbb{E}\left[\left(a_M - \bar{a}_M\right)\Psi_M(\omega)\right] + \mathrm{i}\,B\sqrt{N}\mathbb{E}\left[\left(b_M - \bar{b}_M\right)\Psi_M(\omega)\right]$$

- As in [**H-K-L-N-P, T.IT'2008**], we make intensive use of the **integration by parts formula** ($\mathbf{Z} = \mathbf{D}\mathbf{X}\tilde{\mathbf{D}}$, with $\mathbf{D}, \tilde{\mathbf{D}}$ being diagonal)

$$\mathbb{E}\left[Z_{ij}\Gamma\left(\mathbf{Z}\right)\right] = d_i \tilde{d}_j \mathbb{E}\left[\frac{\partial \Gamma\left(\mathbf{Z}\right)}{\partial Z_{ij}^*}\right]$$

  and the **Nash-Poincaré inequality**

$$\operatorname{var}\left(\Gamma\left(\mathbf{Z}\right)\right) \leq \sum_{i=1}^{M}\sum_{j=1}^{N} d_i \tilde{d}_j \mathbb{E}\left[\left|\frac{\partial \Gamma\left(\mathbf{Z}\right)}{\partial Z_{ij}^*}\right|^2 + \left|\frac{\partial \Gamma\left(\mathbf{Z}\right)}{\partial Z_{ij}^*}\right|^2\right]$$

  to compute the **expectation** and **variance controls** for the following quantities:

$$\operatorname{Tr}\left[\boldsymbol{\Theta}\mathbf{Q}_M^k\right]$$

$$\operatorname{Tr}\left[\boldsymbol{\Theta}\mathbf{Q}_M^k \frac{\mathbf{X}\mathbf{Z}_1\mathbf{X}^H}{N}\right]$$

  where $k = 1, 2, 3, 4$ and $\boldsymbol{\Theta} = \mathbf{a}\mathbf{b}^H$ and $\boldsymbol{\Theta} = \frac{1}{N}\mathbf{Z}_2$ ($\mathbf{a}, \mathbf{b}$ unit-norm and $\mathbf{Z}_1, \mathbf{Z}_2$ diagonal with bounded spectral norm)

- Gathering terms together as

$$\frac{\partial}{\partial\omega}\Psi_M(\omega) = -\omega\left(A^2\sigma_{a^2} + AB\sigma_{ab} + BA\sigma_{ba} + B^2\sigma_{b^2}\right)\Psi_M(\omega) + \mathcal{O}\left(N^{-1}\right)$$

along with $\mathbb{E}[a_M] = \bar{a}_M + \mathcal{O}\left(N^{-1}\right)$ and $\mathbb{E}[b_M] = \bar{b}_M + \mathcal{O}\left(N^{-1}\right)$, we get

$$\sqrt{N}\begin{bmatrix} a_M - \bar{a}_M \\ b_M - \bar{b}_M \end{bmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(\mu, \boldsymbol{\Sigma}), \qquad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{a^2} & \sigma_{ba} \\ \sigma_{ab} & \sigma_{b^2} \end{bmatrix}$$

where $\mu = \mathbf{0}$ and $\sigma_{ab} = \sigma_{ba}$

- Since $\text{SINR}\left(\hat{\mathbf{w}}_{\text{MVDR}}\right) = f(a_M, b_M)$ with $f(x,y) = x^2/y$ and $\nabla f = \begin{bmatrix} 2x/y & -(x/y)^2 \end{bmatrix}$, then it follows by the **Delta method** that

$$\sqrt{N}\left(f(a_M, b_M) - f\left(\bar{a}_M, \bar{b}_M\right)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mu^H \nabla f\left(\bar{a}_M, \bar{b}_M\right), \nabla f\left(\bar{a}_M, \bar{b}_M\right)^H \boldsymbol{\Sigma} \nabla f\left(\bar{a}_M, \bar{b}_M\right)\right)$$

- From the previous procedure we obtain

$$\sigma_M^{-1}\sqrt{N}\left(\mathsf{SINR}\left(\hat{\mathbf{w}}_{\mathsf{MVDR}}\right)-\overline{\mathsf{SINR}\left(\hat{\mathbf{w}}_{\mathsf{MVDR}}\right)}\right)\xrightarrow{\mathcal{L}}\mathcal{N}\left(0,1\right)$$

where

$$\frac{\left(\mathbf{u}^H\mathbf{E}^2\mathbf{u}\right)^2}{\left(\mathbf{u}^H\mathbf{E}\mathbf{u}\right)^4}\sigma_M^2=4\tilde{\gamma}\left(1-\gamma\tilde{\gamma}\right)\mathcal{V}_1$$
$$+4\left(\tilde{\gamma}^2\frac{1}{N}\operatorname{tr}\left[\mathbf{E}^3\right]-\gamma\frac{1}{N}\operatorname{tr}\left[\tilde{\mathbf{E}}^3\right]\right)\mathcal{V}_2+\left(\tilde{\gamma}^2\frac{1}{N}\operatorname{tr}\left[\mathbf{E}^4\right]+\gamma^2\frac{1}{N}\operatorname{tr}\left[\tilde{\mathbf{E}}^4\right]\right)$$
$$+\frac{2}{\left(1-\gamma\tilde{\gamma}\right)}\left(\tilde{\gamma}^3\left(\frac{1}{N}\operatorname{tr}\left[\mathbf{E}^3\right]\right)^2-2\gamma\tilde{\gamma}\frac{1}{N}\operatorname{tr}\left[\mathbf{E}^3\right]\frac{1}{N}\operatorname{tr}\left[\tilde{\mathbf{E}}^3\right]+\gamma^3\left(\frac{1}{N}\operatorname{tr}\left[\tilde{\mathbf{E}}^3\right]\right)^2\right)$$

with

$$\mathcal{V}_1=\left[\frac{\left(\mathbf{u}^H\mathbf{E}^2\mathbf{u}\right)^2}{\left(\mathbf{u}^H\mathbf{E}\mathbf{u}\right)^2}-4\frac{\mathbf{u}^H\mathbf{E}^3\mathbf{u}}{\mathbf{u}^H\mathbf{E}\mathbf{u}}+\frac{1}{2}\left(\frac{\mathbf{u}^H\mathbf{E}^4\mathbf{u}}{\mathbf{u}^H\mathbf{E}^2\mathbf{u}}+\frac{\left(\mathbf{u}^H\mathbf{E}^3\mathbf{u}\right)^2}{\left(\mathbf{u}^H\mathbf{E}^2\mathbf{u}\right)^2}\right)\right]$$
$$\mathcal{V}_2=\left[\frac{\mathbf{u}^H\mathbf{E}^3\mathbf{u}}{\mathbf{u}^H\mathbf{E}^2\mathbf{u}}-\frac{\mathbf{u}^H\mathbf{E}^2\mathbf{u}}{\mathbf{u}^H\mathbf{E}\mathbf{u}}\right]$$

- In the case $\mathbf{T} = \mathbf{I}_N$ and $\alpha = 0$, then $(c = c_M)$

$$\frac{\sqrt{N}}{\|\mathbf{u}\|^2 \sqrt{c\,(1-c)}} \left(\mathsf{SINR}\left(\hat{\mathbf{w}}_{\mathsf{MVDR}}\right) - (1-c)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\,(0,1)$$

- This follows from the CLT-based **Gaussian approximation of the Beta distribution** of $\mathsf{SINR}\left(\hat{\mathbf{w}}_{\mathsf{MVDR}}\right)$ in the finite case by letting $N = N\,(M) \to \infty$

- In the case $\alpha = 0$, then

$$\frac{\sqrt{N}}{\|\mathbf{u}\|^2 \sigma_M}\left(\mathsf{SINR}\left(\hat{\mathbf{w}}_{\mathsf{MVDR}}\right) - \left(1 - \frac{c}{N}\operatorname{Tr}\left[\bar{\mathbf{E}}^2\right]\right)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0,1\right)$$

where

$$\sigma_M^2 = \frac{c^2}{N}\operatorname{Tr}\left[\bar{\mathbf{E}}^4\right] + \frac{c\left(\frac{1}{N}\operatorname{Tr}\left[\bar{\mathbf{E}}^2\right]\right)^2}{\left(1-\frac{c}{N}\operatorname{Tr}\left[\bar{\mathbf{E}}^2\right]\right)} + \frac{c^2\left(\frac{1}{N}\operatorname{Tr}\left[\bar{\mathbf{E}}^2\right]\right)^3 - 4\frac{c}{N}\operatorname{Tr}\left[\bar{\mathbf{E}}^2\right]\frac{c}{N}\operatorname{Tr}\left[\bar{\mathbf{E}}^3\right] + 2c\left(\frac{c}{N}\operatorname{Tr}\left[\bar{\mathbf{E}}^3\right]\right)^2}{\left(1-\frac{c}{N}\operatorname{Tr}\left[\bar{\mathbf{E}}^2\right]\right)}$$

and we have defined $\bar{\mathbf{E}} = \mathbf{T}\left(x_M \mathbf{I}_N + c\mathbf{T}\right)^{-1}$

- We have shown that the SINR of the diagonally loaded Capon/MVDR beamformer is **asymptotically Gaussian** and have provided a closed-form expression for its variance
- The same elements describe also the fluctuations of the MSE performance of this filter, which can be written in terms of **realized variance and bias**, as well as of other linear filters, such as the **linear MMSE filter**
- The results hold for Gaussian environments, but extensions based on a more general integration by parts formula might be investigated for **non-Gaussian observations**
- Rather than on the covariance matrix estimation error, we could directly focused on the performance of the objective by considering the structure of the problem
- A similar scheme can be applied to study the second-order behavior of alternative error measures for **covariance matrix estimation**