E15: Extensions for Complex Data and Parameters

In some applications, models that rely on complex numbers have been found to be useful for various reasons.

Ex) Complex Envelope for Sinusoids

$$s(t) = \sum_{i=1}^{p} A_i \cos(2\pi F_i t + \phi_i) \quad \text{where } F_0 - \frac{B}{2} \le F_i \le F_0 + \frac{B}{2} \quad \forall i.$$

Note that $s(t) = Re\left\{ \sum_{i=1}^{p} A_i e^{j(2\pi F_i t + \phi_i)} \right\}$

$$= 2 Re\left\{ \sum_{i=1}^{p} \frac{A_i}{2} e^{j(2\pi(F_i - F_0)t + \phi_i)} e^{j 2\pi F_0 t} \right\}$$

The Complex Envelope

$$\text{so} \quad \tilde{s}(t) = \sum_{i=1}^{p} \underbrace{\frac{A_i}{2} e^{j\phi_i}}_{\substack{\text{complex} \\ \text{amplitudes}}} e^{j 2\pi \overbrace{(F_i - F_0)}^{}t}$$

frequencies

Sample at the Nyquist rate $F_s = B = 1/\Delta$ ...

$$\tilde{s}(n\Delta) = \sum_{i=1}^{p} \frac{A_i}{2} e^{j\phi_i} e^{j 2\pi (F_i - F_0) n\Delta}$$

$$\Rightarrow \tilde{s}[n] = \sum_{i=1}^{p} \tilde{A}_i e^{j 2\pi f_i n} \quad \text{where } \tilde{A}_i = \frac{A_i}{2} e^{j\phi_i}, \quad f_i = (F_i - F_0)\Delta$$

Data model for signal in additive (complex) noise:

$$\tilde{x}[n] = \tilde{s}[n] + \tilde{w}[n].$$

Ex] Least Squares Estimation of Amplitude

Suppose we wish to minimize $J(\tilde{A}) = \sum_{n=0}^{N-1} |\tilde{x}[n] - \tilde{A}\tilde{s}[n]|^2$ w.r.t. $\tilde{A}$ where all signals and parameters are complex valued. We could express everything in Cartesian form, take the gradient w.r.t. real and imaginary parts of $\tilde{A}$, and equate to zero to find both components of $\tilde{A}$. Alternatively, considering $\hat{A} = \hat{A}_R + j\hat{A}_i$ for this is a complex-valued parameter estimate, we could take the complex derivative and equate to zero.

$$\frac{\partial J}{\partial \tilde{A}} = \frac{1}{2}\left(\frac{\partial J}{\partial A_R} - j\frac{\partial J}{\partial A_i}\right)$$

Note that $\frac{\partial J}{\partial \tilde{A}} = 0$ iff $\frac{\partial J}{\partial \mathbf{A}} = 0$ $\leftarrow \begin{bmatrix} \hat{A}_R \\ \hat{A}_i \end{bmatrix}$

Side note: Let $\theta = \alpha + j\beta$. Consider

$$\frac{\partial \theta}{\partial \theta} = \frac{1}{2}\left(\frac{\partial}{\partial \alpha} - j\frac{\partial}{\partial \beta}\right)(\alpha + j\beta)$$ $\leftarrow$ Definition of complex derivative

$$= \frac{1}{2}\left[\frac{\partial \alpha}{\partial \alpha} + j\frac{\partial \beta}{\partial \alpha} - j\frac{\partial \alpha}{\partial \beta} + \frac{\partial \beta}{\partial \beta}\right] = \frac{1}{2}(1 + j0 - j0 + 1)$$

$$= 1$$

Similarly $\frac{\partial \theta^*}{\partial \theta} = 0$ and $\frac{\partial(\theta\theta^*)}{\partial \theta} = \frac{\partial \theta}{\partial \theta}\theta^* + \theta\frac{\partial \theta^*}{\partial \theta} = \theta^*$

$\therefore \frac{\partial \tilde{A}}{\partial \tilde{A}} = 1$, $\frac{\partial \tilde{A}^*}{\partial \tilde{A}} = 0$, $\frac{\partial(\tilde{A}\tilde{A}^*)}{\partial \tilde{A}} = \tilde{A}^*$, and

$$\frac{\partial J}{\partial \tilde{A}} = \frac{\partial}{\partial \tilde{A}} \sum_{n=0}^{N-1} |\tilde{x}[n] - \tilde{A}\tilde{s}[n]|^2 = \sum_n \Big(|\tilde{x}[n]|^2 - \tilde{x}[n]\tilde{A}^*\tilde{s}^*[n] - \tilde{A}\tilde{s}[n]\tilde{x}^*[n] + \tilde{A}\tilde{A}^*|\tilde{s}[n]|^2 \Big)$$

$$= \sum_{n=0}^{N-1} \left(0 - 0 - \tilde{s}[n]\tilde{x}^*[n] + \tilde{A}^* |\tilde{s}[n]|^2 \right)$$

Setting this equal to zero and solving for $\tilde{A}$ yields the optimal estimate.

Complex Random Variables

Let $\tilde{x} = u + jv$ be a complex-valued random variable. The pdf of $\tilde{x}$ is a joint pdf over $\begin{bmatrix} u \\ v \end{bmatrix}$. We have

$$E[\tilde{x}] = E[u] + jE[v]$$

$$E[|\tilde{x}|^2] = E[u^2] + E[v^2]$$

$$\operatorname{var}(\tilde{x}) = E\left[|\tilde{x} - E[\tilde{x}]|^2 \right] = E[|\tilde{x}|^2] - |E[\tilde{x}]|^2$$

$$\operatorname{cov}(\tilde{x}_1, \tilde{x}_2) = E[\tilde{x}_1^* \tilde{x}_2] - E^*[\tilde{x}_1] E[\tilde{x}_2]$$

$$\tilde{x}_1 \perp\!\!\!\perp \tilde{x}_2 \Rightarrow \operatorname{cov}(\tilde{x}_1, \tilde{x}_2) = 0 \text{ (uncorrelated)}$$

Extension to vectors is straightforward.

$$E[\vec{\tilde{x}}] = \left[E[\tilde{x}_1], \ldots, E[\tilde{x}_n] \right]^T$$

Hermitian $= *T$

$$C_{\vec{\tilde{x}}} = E\left[(\vec{\tilde{x}} - E[\vec{\tilde{x}}])(\vec{\tilde{x}} - E[\vec{\tilde{x}}])^H \right]$$

By construction, $C_{\tilde{x}}^H = C_{\tilde{x}}$ and $C_{\tilde{x}} \geq 0$.

$$\tilde{y} = A\tilde{x} + b \Rightarrow E[\tilde{y}] = AE[\tilde{x}] + b$$

$$C_{\tilde{y}} = AC_{\tilde{x}}A^H$$

complex valued

Let $Q = \tilde{x}^H A \tilde{x}$ where $A^H = A$. Notice that Q is real-valued, since

$$Q^* = (\tilde{x}^H A \tilde{x})^* = \tilde{x}^H A^H \tilde{x} = \tilde{x}^H A \tilde{x} = Q$$

Also if $A > 0$, then $Q > 0 \ \forall \tilde{x} \neq 0$. Assume $E[\tilde{x}] = 0$.

Then $E[Q] = E[\tilde{x}^H A \tilde{x}] = E\{tr(A\tilde{x}\tilde{x}^H)\} = tr(E[A\tilde{x}\tilde{x}^H])$

$$= tr(A\,E[\tilde{x}\tilde{x}^H]) = tr(A\mathbf{C}_{\tilde{x}})$$

$E[Q^2] = E[\tilde{x}^H A \tilde{x}\tilde{x}^H A \tilde{x}]$ will require the fourth order moments of $\tilde{x}$. $E[Q^2] = tr(AC_{\tilde{x}}AC_{\tilde{x}}) + tr^2(AC_{\tilde{x}})$. (p3 513)

Complex Gaussian PDF: $p(\tilde{x}) = \frac{1}{\pi\sigma^2} e^{-\frac{1}{\sigma^2}|\tilde{x}-\tilde{\mu}|^2}$

is obtained directly from $p(u,v) = \frac{1}{\sqrt{2\pi\sigma^2/2}} e^{-\frac{1}{2(\frac{\sigma^2}{2})}(u-\mu_u)^2} \cdot \frac{1}{\sqrt{2\pi\sigma^2/2}} \cdot e^{-\frac{1}{2(\sigma^2/2)}(v-\mu_v)^2}$

by letting $\tilde{\mu} = \mu_u + j\mu_v = E[\tilde{x}]$.

The $p(\tilde{x})$ above is denoted by $CN(\tilde{\mu}, \sigma^2)$ and the complex variance is split equally between the real and imaginary parts (circular symmetry in complex plane).

Assume $\{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n\}$ are independent with $CN(\tilde{\mu}_i, \tilde{\sigma}_i^2)$ for $i = 1, \ldots, n$. Then $p(\vec{\tilde{x}}) = \prod_{i=1}^{n} p(\tilde{x}_i)$ as usual.

Also, since covariances are zero, $C_{\tilde{x}} = diag(\sigma_1^2, \ldots, \sigma_n^2)$.

$$\Rightarrow p(\tilde{x}) = \frac{1}{\pi^n \det(C_{\tilde{x}})} e^{-(\tilde{x}-\tilde{\mu})^H C_{\tilde{x}}^{-1}(\tilde{x}-\tilde{\mu})}$$

also is the (multivariate complex Gaussian)

Thm 15.1 Complex Multivariate Gaussian PDF.

If a real random vector x of dimension $2n \times 1$ can be partitioned as $x = \begin{bmatrix} u \\ v \end{bmatrix}$ where $u \in \mathbb{R}^n$, $v \in \mathbb{R}^n$ and x has the pdf $x \sim \mathcal{N}\left(\begin{bmatrix} \mu_u \\ \mu_v \end{bmatrix}, \begin{bmatrix} C_{uu} & C_{uv} \\ C_{vu} & C_{vv} \end{bmatrix} \right)$

and $C_{uu} = C_{vv}$ and $C_{uv} = -C_{vu}$, then defining the $n \times 1$ complex random vector $\tilde{x} = u + jv$, $\tilde{x}$ has the complex multivariate Gaussian pdf $\tilde{x} \sim \mathcal{CN}(\tilde{\mu}, C_{\tilde{x}})$ where

$\tilde{\mu} = \mu_u + j\mu_v$, $C_{\tilde{x}} = 2(C_{uu} + j\,C_{vu})$, hence

$$p(\tilde{x}) = \frac{1}{\pi^n \det(C_{\tilde{x}})} e^{-(\tilde{x}-\tilde{\mu})^H C_{\tilde{x}}^{-1} (\tilde{x}-\tilde{\mu})}$$

Note that this theorem is made possible because of the isomorphism between 2×2 matrices of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ and complex numbers of the form $a+jb$. Addition and multiplication of ~~[illegible]~~ matrices and complex numbers provide equivalent results.

Many properties of Gaussian pdf's are preserved in the definition for $\mathcal{CN}$, for instance

Gaussian uncorrelated $\Rightarrow$ independent

$\tilde{x} \sim \mathcal{CN}$ & $\tilde{y} = A\tilde{x} + b \Rightarrow \tilde{y} \sim \mathcal{CN}$

Sum of indep. $\mathcal{CN}$ is also $\mathcal{CN}$

$$\begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} \sim \mathcal{CN}\left(\begin{bmatrix} E[\tilde{x}] \\ E[\tilde{y}] \end{bmatrix}, \begin{bmatrix} C_{\tilde{x}\tilde{x}} & C_{\tilde{x}\tilde{y}} \\ C_{\tilde{y}\tilde{x}} & C_{\tilde{y}\tilde{y}} \end{bmatrix} \right)$$

$$\Rightarrow \quad E[\tilde{y} \mid \tilde{x}] = E[\tilde{y}] + C_{\tilde{y}\tilde{x}} C_{\tilde{x}\tilde{x}}^{-1} (\tilde{x} - E[\tilde{x}])$$

$$C_{\tilde{y}|\tilde{x}} = C_{\tilde{y}\tilde{y}} - C_{\tilde{y}\tilde{x}} C_{\tilde{x}\tilde{x}}^{-1} C_{\tilde{x}\tilde{y}}$$

and $p(\tilde{y} \mid \tilde{x}) \sim \mathcal{CN}(E[\tilde{y} \mid \tilde{x}], C_{\tilde{y}|\tilde{x}})$.

Complex WSS Random Process

Consider a complex random process $x[n] = u[n] + jv[n]$.

$$\begin{cases} E[\tilde{x}[n]] = E[u[n]] + jE[v[n]] = \mu_{\tilde{x}} \\ r_{\tilde{x}\tilde{x}}[k] = E[\tilde{x}[n+k]\, \tilde{x}^*[n]] \end{cases}$$

lagging index has $*$ ...

$\Leftrightarrow$ $\tilde{x}[n]$ is WSS

$$\mathcal{F}\{r_{\tilde{x}\tilde{x}}[k]\} = PSD_{\tilde{x}}(w)$$

Notice that
$$\begin{aligned} r_{\tilde{x}\tilde{x}}[k] &= E\left[(u[n+k] + jv[n+k])(u[n] - jv[n])\right] \\ &= E[u[n+k]u[n]] + E[v[n+k]v[n]] \\ &\quad + jE[v[n+k]u[n]] - jE[u[n+k]v[n]] \\ &= r_{uu}[k] + r_{vv}[k] + j r_{uv}[k] - j r_{vu}[k] \end{aligned}$$

we must have $r_{uu}[k] = r_{vv}[k]$ and $r_{uv}[k] = -r_{vu}[k]$ when $\tilde{x}[n]$ is a WSS Gaussian random process, as a generalization of $C_{uu} = C_{vv}$ and $C_{uv} = -C_{vu}$.

$\therefore$ $P_{uu}(f) = P_{vv}(f)$ and $P_{uv}(f) = -P_{vu}(f)$.

In other words, $r_{\tilde{x}\tilde{x}}[k] = 2r_{uu}[k] + 2j\, r_{uv}[k]$

$$P_{\tilde{x}\tilde{x}}(f) = 2\left(P_{uu}(f) + j P_{uv}(f)\right)$$

Derivatives and Optimization

Let $\theta = \alpha + j\beta$, $J(\theta) \in \mathbb{R}$, $\frac{\partial J}{\partial \theta} = \frac{1}{2}\left(\frac{\partial J}{\partial \alpha} - j\frac{\partial J}{\partial \beta}\right)$.

When $\boldsymbol{\theta} \in \mathbb{C}^p$, $\frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \left[\frac{\partial J}{\partial \theta_1} \cdots \frac{\partial J}{\partial \theta_p}\right]^T$.

The following are useful:

$$\frac{\partial \theta}{\partial \theta} = 1, \quad \frac{\partial \theta^*}{\partial \theta} = 0 \qquad \frac{\partial b^H \theta}{\partial \theta} = b^*, \quad \frac{\partial \theta^H b}{\partial \theta} = 0$$

for $\xi_i \in \mathbb{R}$, $\frac{\partial \ln\det(C_{\tilde{x}}(\xi))}{\partial \xi_i} = \mathrm{tr}\left(C_{\tilde{x}}^{-1}(\xi)\frac{\partial C_{\tilde{x}}(\xi)}{\partial \xi_i}\right)$ $\quad \frac{\partial \theta^H A \theta}{\partial \theta} = (A\theta)^*$, where $A^H = A$

$$\frac{\partial \tilde{x}^H C_{\tilde{x}}^{-1}(\xi)\tilde{x}}{\partial \xi_i} = -\tilde{x}^H C_{\tilde{x}}^{-1}(\xi)\frac{\partial C_{\tilde{x}}(\xi)}{\partial \xi_i} C_{\tilde{x}}^{-1}(\xi)\tilde{x}$$

Ex| Minimization of Hermitian Functions

$$J = (\tilde{x} - H\theta)^H C^{-1}(\tilde{x} - H\theta) \qquad \text{(complex least squares)}$$

where $C^H = C$. Note that J is real $(J^H = J^* = J)$

$$J = \tilde{x}^H C^{-1}\tilde{x} - \tilde{x}^H C^{-1} H\theta - \theta^H H^H C^{-1}\tilde{x} + \theta^H H^H C^{-1} H\theta$$

$$\frac{\partial J}{\partial \theta} = 0 - (H^H C^{-1}\tilde{x})^* - 0 + (H^H C^{-1} H\theta)^* = -\left[H^H C^{-1}(\tilde{x} - H\theta)\right]^*$$

Equating to zero and solving $\Rightarrow \hat{\theta} = (H^H C^{-1} H)^{-1} H^H C^{-1}\tilde{x}$

We can verify that $\hat{\theta}$ is a global minimum by showing that $J(\hat{\theta}) \geq J(\theta)\ \forall \theta$. Showing further that equality is only achieved at $\hat{\theta}$ verifies that it is the only global minimizer.

Ex) Minimization of a Hermitian form subject to linear constraints.

$$\min_{a}\ a^H W a \quad \text{s.t.}\quad Ba = b \qquad a \in \mathbb{C}^p,\ b \in \mathbb{C}^r_{\ r<p}$$

$$B = B_R + jB_I,\quad a = a_R + ja_I,\quad b = b_R + jb_I$$

$$\Rightarrow (B_R + jB_I)(a_R + ja_I) = b_R + jb_I$$

$$\Leftrightarrow \quad B_R a_R - B_I a_I = b_R$$

$$B_I a_R + B_R a_I = b_I$$

Introducing Lagrange multipliers

$$J(a) = a^H W a + \lambda_R^T (B_R a_R - B_I a_I - b_R) + \lambda_I^T (B_I a_R + B_R a_I - b_I)$$

$$= a^H W a + \lambda_R^T \operatorname{Re}\{Ba - b\} + \lambda_I^T \operatorname{Im}\{Ba - b\}$$

$$= a^H W a + \operatorname{Re}\{(\lambda_R - j\lambda_I)^T (Ba - b)\}$$

with $\lambda = \lambda_R + j\lambda_I$

$$= a^H W a + \frac{1}{2}\lambda^H (Ba - b) + \frac{1}{2}\lambda^T (B^* a^* - b^*)$$

Then $\frac{\partial J}{\partial a} = (Wa)^* + \left(\frac{B^H \lambda}{2}\right)^*$ and setting it to zero yields

$$a_{opt} = -W^{-1} B^H \frac{\lambda}{2}$$

Imposing the constraint $B a_{opt} = b$, we get

$$\frac{\lambda}{2} = -(B W^{-1} B^H)^{-1} b$$

$$\Rightarrow a_{opt} = W^{-1} B^H (B W^{-1} B^H)^{-1} b$$

We can show that this is the global minimizer by verifying $a^H W a \geq a_{opt}^H W a_{opt}$ and equality is attained only for $a = a_{opt}$.

Classical Estimation with Complex Data

Consider complex Gaussian data for illustrative purposes.

$$p(\tilde{x}, \theta) = \frac{1}{\pi^N \det C_{\tilde{x}}(\theta)} e^{-(\tilde{x} - \tilde{\mu}(\theta))^H C_{\tilde{x}}^{-1}(\theta)(\tilde{x} - \tilde{\mu}(\theta))}$$

where $\tilde{x} = [\tilde{x}[0], \tilde{x}[1], \ldots, \tilde{x}[N-1]]^T$ and $\theta \in \mathbb{C}^{P/2}$ ~~most~~ if all params are complex.

We will use $\xi \in \mathbb{R}^P$ to denote the vector of real-valued parameters

Fisher Information Matrix:

$$[I(\xi)]_{ij} = \operatorname{tr}\left[C_{\tilde{x}}^{-1}(\xi) \frac{\partial C_{\tilde{x}}(\xi)}{\partial \xi_i} C_{\tilde{x}}^{-1}(\xi) \frac{\partial C_{\tilde{x}}(\xi)}{\partial \xi_j} \right] + 2 \operatorname{Re}\left[\frac{\partial \tilde{\mu}^H(\xi)}{\partial \xi_i} C_{\tilde{x}}^{-1}(\xi) \frac{\partial \tilde{\mu}(\xi)}{\partial \xi_j} \right]$$

for $i, j = 1, \ldots, P$ (derivation in App 15C).

The equality is attained if

$$\frac{\partial \ln p(\tilde{x};\xi)}{\partial \xi} = I(\xi)\left(g(\tilde{x})-\xi\right)$$

where $g(\tilde{x})$ is the efficient estimator for ξ.

Ex) Complex Classical Linear Model

Assume $\tilde{x} = H\theta + \tilde{w}$ where H is a known complex $N \times p$ matrix where $N > p$, H is full rank, θ is complex $p \times 1$, $\tilde{w}$ is complex $N \times 1$ with $\tilde{w} \sim \mathcal{CN}(0, C)$. Then $\tilde{x} \sim \mathcal{CN}(H\theta, C)$.

Let $\tilde{\mu} = H\theta$, $C(\theta) = C$. We have

$$p(\tilde{x};\theta) = \frac{1}{\pi^N \det(C)} e^{-(\tilde{x}-H\theta)^H C^{-1}(\tilde{x}-H\theta)}$$

$$\frac{\partial \ln p(\tilde{x};\theta)}{\partial \theta^*} = H^H C^{-1}(\tilde{x}-H\theta)$$

$$= \underbrace{H^H C^{-1} H}_{I(\theta)}\Big[\underbrace{(H^H C^{-1} H)^{-1} H^H C^{-1}\tilde{x}}_{g(\tilde{x})} - \theta\Big]$$

The equality condition is satisfied and

$$\hat{\theta} = (H^H C^{-1} H)^{-1} H^H C^{-1}\tilde{x} \quad \text{with } C_{\hat{\theta}} = (H^H C^{-1} H)^{-1}$$

is an efficient estimator (and hence the MVU estimator).

Bayesian Estimation

If $\tilde{x}$ and θ are jointly Gaussian, then the posterior pdf $p(\theta|\tilde{x}) \sim CN$ with

$$E[\theta|\tilde{x}] = E[\theta] + C_{\theta\tilde{x}} C_{\tilde{x}\tilde{x}}^{-1} (\tilde{x} - E[\tilde{x}])$$

$$C_{\tilde{\theta}|\tilde{x}} = C_{\theta\theta} - C_{\theta\tilde{x}} C_{\tilde{x}\tilde{x}}^{-1} C_{\tilde{x}\theta}$$

In the CN case, MMSE and MAP estimators are identical. $B_{MSE}(\hat{\theta}_i) = E[|\theta_i - \hat{\theta}_i|^2]$

Ex) Bayesian Linear Model $(\tilde{x} = H\theta + \tilde{w})$

Assume $\theta \sim CN(\mu_\theta, C_{\theta\theta})$. $E[\tilde{x}] = H\mu_\theta$ since $\tilde{w} \sim CN(0, C_{\tilde{w}})$ is indep. of θ.

$$C_{\tilde{x}\tilde{x}} = H C_{\theta\theta} H^H + C_{\tilde{w}} \quad \text{and} \quad C_{\theta\tilde{x}} = C_{\theta\theta} H^H.$$

We also have $\begin{bmatrix} \tilde{x} \\ \theta \end{bmatrix} = \begin{bmatrix} H & I \\ I & 0 \end{bmatrix} \begin{bmatrix} \theta \\ \tilde{w} \end{bmatrix}$, so $\begin{bmatrix} \tilde{x} \\ \theta \end{bmatrix}$ is jointly (complex) Gaussian. Then

$$\hat{\theta} = \mu_\theta + C_{\theta\theta} H^H (H C_{\theta\theta} H^H + C_{\tilde{w}})^{-1} (\tilde{x} - H\mu_\theta)$$

$$= \mu_\theta + (C_{\theta\theta}^{-1} + H^H C_{\tilde{w}}^{-1} H)^{-1} H^H C_{\tilde{w}}^{-1} (x - H\mu_\theta)$$

and $B_{MSE}(\hat{\theta}_i) = [C_{\theta\theta} - C_{\theta\theta} H^H (H C_{\theta\theta} H^H + C_{\tilde{w}})^{-1} H C_{\theta\theta}]_{ii}$

$$= [(C_{\theta\theta}^{-1} + H^H C_{\tilde{w}}^{-1} H)^{-1}]_{ii}$$

Suggested Problems: 5, 10, 16, 18, 24