ERRATA AND COMMENTS FOR ESE 524 SLIDES

Spring 2018

Errata and Comments for l0.pdf

1. All references to EE 527 at the bottom of the slides should be replaced with ESE 524.

2. Page 2, 17, 18, 19, and 20, skip these pages.

3. Page 6, if in our data model, θ is an unknown deterministic parameter, then we usually write the model into $p(x;\theta)$. If we model the parameter θ as a random variable, then we usually write the data model into $p(x|\theta)$. However, in some textbooks or papers, authors do not distinguish between these notations.

4. Page 8, the example given for the online estimator on this page is not clear. Here is a better example ($\hat{\theta} = \hat{s}[n]$). Define

$$\hat{s}[n] = a_0 x[0] + a_1 x[1] + \ldots + a_n x[n], \;\; n = 0, 1, 2, \ldots$$

   Thus, we can see its online mode:

$$\hat{s}[n] = \hat{s}[n-1] + a_n x[n]$$

5. Page 9, usually $p(\boldsymbol{x}|\boldsymbol{\theta})$ is written as $p(\boldsymbol{x};\boldsymbol{\theta})$ for the ML method, since $\boldsymbol{\theta}$ is an unknown deterministic parameter rather than a random variable.

6. Page 10, the form of MMSE in the second bullet in proved on page 49-51 in l4.pdf.

7. Page 11, in the slides, we interchangeably use classical approach and non-Bayesian approach. The non-Bayesian approach is also called frequentist approach in some literature.

8. Page 12, skip: "Notes: we do not care much about identifiability ..."

9. Page 14, (2) is shown on page 23-24 in l3.pdf.

Errata and Comments for l1.pdf

1. Page 1, skip: "Go over handouts 2–5 in EE 420x notes".

2. Page 2, proof of (5):
   Equation (5) describes the *law of total probability*. Because B_1, B_2, $\ldots$, B_n form a *partition* of Ω, the following holds:

   (1) $B_i \cap B_j = \emptyset$, $\forall i \neq j$;

   (2) $B_1 \cup B_2 \cup \cdots \cup B_n = \Omega$.

   Hence

$$\begin{aligned}
\mathrm{P}[A] &= \mathrm{P}[A \cap \Omega] \\
&= \mathrm{P}[A \cap B_1 + A \cap B_2 + \cdots + A \cap B_n] \\
&= \mathrm{P}[A \cap B_1] + \cdots + \mathrm{P}[A \cap B_n] \\
&= \mathrm{P}[A|B_1]P[B_1] + \cdots + \mathrm{P}[A|B_n]\mathrm{P}[B_n].
\end{aligned}$$

   The last line follows directly from (4) on page 1.

3. Page 3, skip: "(To refresh memory about ...)".

4. Page 3, skip: "(handout 5 in EE 420x notes)".

5. Page 4, proof of (2'):
   To prove this, we make use of the properties in (1) on this page, as well as the definition of $\mathrm{cov}(\boldsymbol{X})$. To make the derivations clearer, for every expectation operation, we explicitly write down the random variable with respect to which the expectation is taken.

$$\begin{aligned}
\mathrm{cov}[\boldsymbol{X}] &= \mathrm{E}_{\boldsymbol{X}}[\boldsymbol{X}\boldsymbol{X}^T] - \mathrm{E}_{\boldsymbol{X}}[\boldsymbol{X}]\mathrm{E}_{\boldsymbol{X}}[\boldsymbol{X}]^T \\
&= \mathrm{E}_{\boldsymbol{Y}}\big[\mathrm{E}_{\boldsymbol{X}|\boldsymbol{Y}}[\boldsymbol{X}\boldsymbol{X}^T|\boldsymbol{Y}]\big] - \mathrm{E}_{\boldsymbol{Y}}\big[\mathrm{E}_{\boldsymbol{X}|\boldsymbol{Y}}[\boldsymbol{X}|\boldsymbol{Y}]\big]\mathrm{E}_{\boldsymbol{Y}}\big[\mathrm{E}_{\boldsymbol{X}|\boldsymbol{Y}}[\boldsymbol{X}|\boldsymbol{Y}]\big]^T \\
&= \mathrm{E}_{\boldsymbol{Y}}\big[\mathrm{cov}(\boldsymbol{X}|\boldsymbol{Y}) + \mathrm{E}_{\boldsymbol{X}|\boldsymbol{Y}}[\boldsymbol{X}|\boldsymbol{Y}]\mathrm{E}_{\boldsymbol{X}|\boldsymbol{Y}}[\boldsymbol{X}|\boldsymbol{Y}]^T\big] \\
&\quad - \mathrm{E}_{\boldsymbol{Y}}\big[\mathrm{E}_{\boldsymbol{X}|\boldsymbol{Y}}[\boldsymbol{X}|\boldsymbol{Y}]\big]\mathrm{E}_{\boldsymbol{Y}}\big[\mathrm{E}_{\boldsymbol{X}|\boldsymbol{Y}}[\boldsymbol{X}|\boldsymbol{Y}]\big]^T \\
&= \mathrm{E}_{\boldsymbol{Y}}[\mathrm{cov}(\boldsymbol{X}|\boldsymbol{Y})] + \big(\mathrm{E}_{\boldsymbol{Y}}\big[\mathrm{E}_{\boldsymbol{X}|\boldsymbol{Y}}[\boldsymbol{X}|\boldsymbol{Y}]\mathrm{E}_{\boldsymbol{X}|\boldsymbol{Y}}[\boldsymbol{X}|\boldsymbol{Y}]^T\big] \\
&\quad - \mathrm{E}_{\boldsymbol{Y}}\big[\mathrm{E}_{\boldsymbol{X}|\boldsymbol{Y}}[\boldsymbol{X}|\boldsymbol{Y}]\big]\mathrm{E}_{\boldsymbol{Y}}\big[\mathrm{E}_{\boldsymbol{X}|\boldsymbol{Y}}[\boldsymbol{X}|\boldsymbol{Y}]\big]^T\big) \\
&= \mathrm{E}_{\boldsymbol{Y}}[\mathrm{cov}(\boldsymbol{X}|\boldsymbol{Y})] + \mathrm{cov}(\mathrm{E}_{\boldsymbol{X}|\boldsymbol{Y}}[\boldsymbol{X}|\boldsymbol{Y}])
\end{aligned}$$

6. Page 4, typo: that $h(\cdot)$ is the unique inverse of $g(\cdot)$ should be that $\boldsymbol{h}(\cdot)$ is the unique inverse of $\boldsymbol{g}(\cdot)$. Note that

$$\boldsymbol{Y} = \boldsymbol{g}(\boldsymbol{X}) \;\Leftrightarrow\; \boldsymbol{h}(\boldsymbol{Y}) = \boldsymbol{X}.$$

   Also, $p_X(h_1(y_1), \ldots, h_n(y_n))$ should be $p_X(h_1(\boldsymbol{y}), \ldots, h_n(\boldsymbol{y})) = p_X(\boldsymbol{h}(\boldsymbol{y}))$, where $\boldsymbol{h}(\cdot) = [h_1(\cdot), \ldots, h_n(\cdot)]$. Further, $\frac{\partial x_i}{\partial y_j} = \frac{\partial h_i(\boldsymbol{y})}{\partial y_j}$ for the Jacobian determinant J. A good reference, which gives a correct version, is on page 7 in lecture notes 02_revprob.pdf.

7. Page 4, skip: "Print and read the handout "Probability distributions" from the Supplementary material section on WebCT. Bring it with you to the midterm exam."

8. Page 5, cross reference for "Note that $\hat{A}_1$ is the ML ...": page 14 in l0.pdf.

9. Page 6, typo: in the last equation, $w[n]$ should be $w[0]$.

10. Page 8, Bayesians' criticism can be seen on page 11 in l0.pdf.

11. Page 17, skip: "Reading: Kay-I, chs. 5.3–5.4."

12. Page 19, typo: $\sum_{t=1}^{n} x_i$ should be $\sum_{i=1}^{n} x_i$.

13. Page 24, "similar to the one in handout # 0" refers to the example on page 13 in l0.pdf.

14. Page 34, (7) can be proved by factorization theorem. Specifically, we have

$$\begin{aligned} p(\boldsymbol{x};\boldsymbol{\alpha}) &= \frac{1}{\sqrt{(2\pi)^n|\boldsymbol{C}|}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{F\alpha})^T\boldsymbol{C}^{-1}(\boldsymbol{x}-\boldsymbol{F\alpha})\right\} \\ &= \frac{1}{\sqrt{(2\pi)^n|\boldsymbol{C}|}} \exp\left\{-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{C}^{-1}\boldsymbol{x}+\frac{1}{2}\boldsymbol{\alpha}^T\boldsymbol{F}^T\boldsymbol{C}^{-1}\boldsymbol{x}+\frac{1}{2}\boldsymbol{x}^T\boldsymbol{C}^{-1}\boldsymbol{F\alpha}-\frac{1}{2}\boldsymbol{\alpha}^T\boldsymbol{F}\boldsymbol{C}^{-1}\boldsymbol{F\alpha}\right\} \\ &= \underbrace{\frac{1}{\sqrt{(2\pi)^n|\boldsymbol{C}|}} \exp\left\{\boldsymbol{\alpha}^T\boldsymbol{F}^T\boldsymbol{C}^{-1}\boldsymbol{x}-\frac{1}{2}\boldsymbol{\alpha}^T\boldsymbol{F}\boldsymbol{C}^{-1}\boldsymbol{F\alpha}\right\}}_{g(T(\boldsymbol{x}),\boldsymbol{\alpha})} \cdot \underbrace{\exp\left\{-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{C}^{-1}\boldsymbol{x}\right\}}_{h(\boldsymbol{x})} \end{aligned}$$

where $T(\boldsymbol{x}) = \boldsymbol{F}^T\boldsymbol{C}^{-1}\boldsymbol{x}$. Thus, by factorization theorem, $T(\boldsymbol{x}) = \boldsymbol{F}^T\boldsymbol{C}^{-1}\boldsymbol{x}$ is a sufficient statistic for $\boldsymbol{\alpha}$, when $\boldsymbol{C}$ is known.

15. Page 38, skip: "e.g. the EM algorithm (to be discussed later)". The EM algorithm will not be discussed in this course.

Errata and Comments for l2.pdf

1. Page 1, skip: "Reading Kay-I, Ch. 3."
2. Page 9, typo: the second line from the bottom, "that" should be "than".
3. Page 18, typo: AWNG should be AWGN.
4. Page 19, typo: $\mathcal{I}(f)$ should be

$$\mathcal{I}(f) = -\mathrm{E}_{p(\boldsymbol{x};f)}\left[\frac{\partial^2 \log p(\boldsymbol{X};f)}{\partial f^2}\right]$$

5. Page 26, typo: in the second itemized sentence "Suppose now that...", $\psi(\boldsymbol{\theta}) = \theta_i$ should be $\psi_i(\boldsymbol{\theta}) = \theta_i$. Also, $\frac{\partial \psi(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ should be $\frac{\partial \psi_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$.
6. Page 28-30, typos:
   (a) $\mathrm{E}_{p(x;\boldsymbol{\theta})}[T_i(X)]$ should be $\mathrm{E}_{p(x;\boldsymbol{\eta})}[T_i(X)]$ on page 28.
   (b) Two $\mathrm{var}_{p(x;\boldsymbol{\theta})}[T_1(X)]$ should be $\mathrm{var}_{p(x;\boldsymbol{\eta})}[T_1(X)]$ on page 29.
   (c) $\frac{\partial \boldsymbol{A}(\boldsymbol{\eta})}{\partial \eta_i \partial \boldsymbol{\eta}}$ should be $\frac{\partial \boldsymbol{A}(\boldsymbol{\eta})}{\partial \eta_1 \partial \boldsymbol{\eta}}$ on page 29.
   (d) $\mathrm{E}_{p(x;\boldsymbol{\theta})}[T_i(X)]$ should be $\mathrm{E}_{p(x;\boldsymbol{\eta})}[T_i(X)]$ on page 30.
7. Page 30-31, Typo: "the MVU" should be "a MVU".
8. Page 36-45, skip these pages.

Errata and Comments for l3.pdf

1. Page 8, typo: $\boldsymbol{\theta} = \boldsymbol{H}\backslash\boldsymbol{x}$ should be $\hat{\boldsymbol{\theta}} = \boldsymbol{H}\backslash\boldsymbol{x}$.

2. Page 13, skip: "Note the equivalence with MVDR beamforming. To read more about MVDR beamforming, see ...", as MVDR is beyond the scope of this course.

3. Page 17, typo: $\hat{\sigma}^2 = (1/N)\sum_{n=0}^{N-1} x^2[n]$ should be $\hat{\sigma}^2 = (1/N)\sum_{n=1}^{N} x^2[n]$.

4. Page 20, typo: $\hat{\sigma}^2 = \frac{1}{N}\sum_{n=1}^{N}(x[n]-\bar{x})^2$ should be $\hat{\sigma}^2 = \frac{1}{N-1}\sum_{n=1}^{N}(x[n]-\bar{x})^2$

5. Page 26, the concepts "minimal" and "complete" are not well defined in this set of slides.

6. Page 32, typo: in the last sentence, "smaller" should be "larger."

7. Page 52, typo: in the first two equation, $t = 1$ should be $i = 1$.

8. Page 67, derivation for the first two equations:

$$
\begin{aligned}
&||\boldsymbol{x} - \boldsymbol{H}\boldsymbol{\theta}||^2 \\
&\quad = (\boldsymbol{x}^T - \boldsymbol{\theta}^T\boldsymbol{H}^T)\boldsymbol{I}(\boldsymbol{H}\boldsymbol{\theta} - \boldsymbol{x}) \\
&\quad = (\boldsymbol{x}^T - \boldsymbol{\theta}^T\boldsymbol{H}^T)\boldsymbol{Q}\boldsymbol{Q}^T(\boldsymbol{H}\boldsymbol{\theta} - \boldsymbol{x}) \\
&\quad = (\boldsymbol{x}^T\boldsymbol{Q} - \boldsymbol{\theta}^T\boldsymbol{H}^T\boldsymbol{Q})(\boldsymbol{Q}^T\boldsymbol{H}\boldsymbol{\theta} - \boldsymbol{Q}^T\boldsymbol{x}) \\
&\quad = (\boldsymbol{x}^T\boldsymbol{Q} - \boldsymbol{\theta}^T\boldsymbol{R}^T\boldsymbol{Q}^T\boldsymbol{Q})(\boldsymbol{Q}^T\boldsymbol{Q}\boldsymbol{R}\boldsymbol{\theta} - \boldsymbol{Q}^T\boldsymbol{x}) \\
&\quad = (\boldsymbol{x}^T\boldsymbol{Q} - \boldsymbol{\theta}^T\boldsymbol{R}^T)(\boldsymbol{R}\boldsymbol{\theta} - \boldsymbol{Q}^T\boldsymbol{x}) = ||\boldsymbol{Q}^T\boldsymbol{x} - \boldsymbol{R}\boldsymbol{\theta}||^2 \;\; \rightarrow \text{the first equation} \\
&\quad = \left\| \begin{bmatrix} \boldsymbol{Q}_1^T \\ \boldsymbol{Q}_2^T \end{bmatrix} \boldsymbol{x} - \begin{bmatrix} \boldsymbol{R}_1 \\ \boldsymbol{0} \end{bmatrix} \boldsymbol{\theta} \right\|^2 \\
&\quad = \left\| \begin{bmatrix} \boldsymbol{Q}_1^T\boldsymbol{x} - \boldsymbol{R}_1\boldsymbol{\theta} \\ \boldsymbol{Q}_2^T\boldsymbol{x} \end{bmatrix} \right\|^2 \\
&\quad = \begin{bmatrix} \boldsymbol{Q}_1^T\boldsymbol{x} - \boldsymbol{R}_1\boldsymbol{\theta} \\ \boldsymbol{Q}_2^T\boldsymbol{x} \end{bmatrix}^T \begin{bmatrix} \boldsymbol{Q}_1^T\boldsymbol{x} - \boldsymbol{R}_1\boldsymbol{\theta} \\ \boldsymbol{Q}_2^T\boldsymbol{x} \end{bmatrix} \\
&\quad = \left\|\boldsymbol{Q}_1^T\boldsymbol{x} - \boldsymbol{R}_1\boldsymbol{\theta}\right\|^2 + \left\|\boldsymbol{Q}_2^T\boldsymbol{x}\right\|^2 \;\; \rightarrow \text{the second equation}
\end{aligned}
$$

Errata and Comments for l4.pdf

1. Page 8, delete "we use (6) and verify its validity." Here, it is meant to verify (6) using this example.

2. Page 9, after "we apply (8)", it would be more logical to add "$p(x_2|x_1) =$" before the integral, i.e.,
$$p(x_2|x_1) = \int_0^1 p(x_2|\theta) \cdot p(\theta|x_1) d\theta.$$

3. Page 16, typo:
$$p(\theta|\boldsymbol{x}) = p(\theta|\bar{x}) \propto \pi(\theta) p(\bar{x}|\theta) = \mathcal{N}(\mu_N, \tau_N^2)$$
should be
$$p(\theta|\boldsymbol{x}) = p(\theta|\bar{x}) \propto \pi(\theta) p(\bar{x}|\theta) \propto \mathcal{N}(\mu_N, \tau_N^2)$$

4. Page 21, add "Gaussian" after i.i.d..

5. Page 22, typo: $p(\boldsymbol{x}|\boldsymbol{\theta})$ should be $p(\boldsymbol{x}|\theta)$ to make the notation consistent.

6. Page 22, typo:
$$\exp\left[-\frac{1}{2\tau_N^2}(\mu_N - \bar{x})\right]$$
should be
$$\exp\left[-\frac{1}{2\tau_N^2}(\mu_N - \theta)\right]$$

7. Page 23, typo:
$$p(\theta|\boldsymbol{x}) \propto [g(\theta)]^{n+\eta} \exp\{\boldsymbol{\phi}(\theta)^T [\boldsymbol{t}(\boldsymbol{x} + \boldsymbol{\nu})]\}$$
should be
$$p(\theta|\boldsymbol{x}) \propto [g(\theta)]^{N+\eta} \exp\{\boldsymbol{\phi}(\theta)^T [\boldsymbol{t}(\boldsymbol{x} + \boldsymbol{\nu})]\}$$

8. Page 27, more precisely, σ_0^4/ν_0 should be $2\sigma_0^4/\nu_0$.

9. Page 37, typo: x_* should be $\boldsymbol{x}_*$.

10. Page 38, delete "These approaches will be discussed in detail later."

11. Page 40, typo: Inv-$\chi^2(N-1, s^2)$ should be Inv-$\chi^2(N-1, \frac{N}{N-1}s^2)$.

12. Page 45, typos: for the red notes in the figure, the left one should be $\frac{1-c}{2}$ quantile, and the right one should be $(1 - \frac{1-c}{2})$ quantile.

13. Page 63, here we also assume that $\boldsymbol{\theta}$ and $\boldsymbol{w}$ are independent.

14. Page 64-67, skip these pages.

15. Page 77-126, skip these pages.

Errata and Comments for 30_Kalman.pdf

1. Page 5, typo: on the right-hand-side of the second equality in (15), $f_{\boldsymbol{B}_{k-1}|\boldsymbol{Y}_{1:(k-1)}}(\boldsymbol{B}|\boldsymbol{y}_{1:(k-1)})$ should be
$$f_{\boldsymbol{B}_{k-1}|\boldsymbol{Y}_{1:(k-1)}}\left(\boldsymbol{\beta}|\boldsymbol{y}_{1:(k-1)}\right)$$

Errata and Comments for l5.pdf

1. Page 14, typo: "Recall that $\Lambda(x)$ is the sufficient ..." should be "Recall that $\Lambda(x)$ is a sufficient ..."

2. Page 23, $1/N$ is missing before $\log(\pi_0^{\lambda}\pi_1^{1-\lambda})$.

3. Page 25, in the definition of P_{D}, it should be $\boldsymbol{X} \in \mathcal{X}_1$.

4. Page 26, in comment (iii), $\mathcal{R}_1$ should be $\mathcal{X}_1$.

5. Page 27, typo:
$$P_{\mathrm{D}} = P[\boldsymbol{X} \in \mathcal{X}_1; \theta = \theta_0]$$
should be
$$P_{\mathrm{D}} = P[\boldsymbol{X} \in \mathcal{X}_1; \theta = \theta_1]$$

6. Page 38, there should be a minus sign before the first $\min_{\lambda\in[0,1]}$. The second $\min_{\lambda\in[0,1]}$ should be $\max_{\lambda\in[0,1]}$.

7. Page 54, $\sum_{i=0}^{M-1}\int_{\Theta_i} p(\boldsymbol{x}|\theta)\pi(\theta)d\theta$ should be defined as $h_m(\boldsymbol{x})$ instead of $h_m(\theta)$. The x should be $\boldsymbol{x}$ in $h_m(\boldsymbol{x})$.

Errata and Comments for Bayesian Detection Examples

1. Page 3, Line 3, "This terms varies ..." should be "This term varies ..."

2. Page 5, equation (5) should be
$$f_{T(\boldsymbol{X}|\Lambda)}(T(\boldsymbol{x})|\lambda) = \frac{1}{T(\boldsymbol{x})!}(N\lambda)^{T(\boldsymbol{X})}e^{-N\lambda}.$$

3. Page 6, the expressions of P_{FA} and P_{D} should be corrected as
$$P_{\mathrm{FA}} = \sum_{m=\lceil N\eta' \rceil}^{+\infty} \frac{(N\lambda_0)^m}{m!}e^{-N\lambda_0},$$
$$P_{\mathrm{D}} = \sum_{m=\lceil N\eta' \rceil}^{+\infty} \frac{(N\lambda_1)^m}{m!}e^{-N\lambda_1}.$$