Homework Assignment – 601.455/655 Fall 2025

Name	Name
Email	Email
Other contact information (optional)	Other contact information (optional)
Signature (required) I/We have followed the rules in completing this assignment ________________________________	Signature (required) I/We have followed the rules in completing this assignment ________________________________

1. Remember that this is a graded homework assignment. It is the functional equivalent of a take-home exam.
2. You are to work **alone** or in **teams of two** and are not to discuss the problems with anyone other than the TAs or the instructor.
3. **IMPORTANT NOTE:** If you work in teams of two, you are **not** to split up the questions and each answer a subset individually. You are to work **together**. I encourage teaming on these problems because I believe that it encourages learning, not as a way to reduce the required work for students taking the course. By signing this sheet you are asserting that each of you has contributed equally to each answer and can individually explain the answer as well as if you had answered the question alone. I view this as a question of trust and ethics.
4. It is otherwise open book, notes, and web. But you should cite any references you consult.
5. Please refer to the course organizational notes for a fuller listing of all the rules. I am not reciting them all here, but they are still in effect.
6. Unless I say otherwise in class, it is due before the start of class on the due date posted on the web.
7. Submit the assignment on GradeScope as a neat and legible PDF file. We will not insist on typesetting your answers, but we must be able to read them. We will not go to extraordinary lengths to decipher what you write. If the graders cannot make out an answer, the score will be 0.
8. Sign and hand in this page as the first sheet of your assignment. If you work with a partner, then you both should sign the sheet, but you should only submit one PDF file for both of you, using the GradeScope teaming feature. Indicate clearly who it is from.
9. This assignment has more than 100 points, but the most that will be applied to your grade is 100.

Notes on this assignment

Notational Notes

Generally, one would represent the covariance of a random variable X using a single subscript $cov(X) = \Sigma_X$ or $cov(X) = \mathbf{C}_X$ and the covariance of two random variables (X,Y) as $\text{cov}(X,Y) = \Sigma_{XY}$ or $\text{cov}(X,Y) = \mathbf{C}_{XY}$. Sometimes I will double the subscript for covariance of a single random variable, e.g., $\text{cov}(X) = \text{cov}(X,X) = \Sigma_{XX} = \mathbf{C}_{XX}$.

The questions below concern iterative updates of the distributions of random variables. For example, $\vec{\mathbf{x}}^{(k)} \sim N\left(\vec{\mu}_{\vec{x}}^{(k)}, \mathbf{C}_{\vec{x}\vec{x}}^{(k)}\right)$ would represent a multivariable Gaussian distribution for the random vector $\vec{\mathbf{x}}$. To reduce clutter, we will sometimes omit the vector symbol in subscripts. E.g., $\vec{\mathbf{x}}^{(k)} \sim N\left(\vec{\mu}_{x}^{(k)}, \mathbf{C}_{xx}^{(k)}\right)$. We will omit the superscript for random variables representing quantities that are unique to an individual observation, and use the superscript (*obs*)" for an observed value. For example, $\vec{\mathbf{y}}_k^{(obs)}$ would represent an observed value of random vector $\vec{\mathbf{y}}_k$ with $\vec{\mathbf{y}}_k \sim N\left(\vec{\mu}_{y,k}, \mathbf{C}_{yy,k}\right)$.

Other Notes

Multivariable Gaussian: Given a *d*-dimensional Gaussian random vector $\vec{\mathbf{x}} \sim N(\mu_x, \mathbf{C}_{xx})$, then the probability that $\vec{\mathbf{x}}$ has any particular value $\vec{\mathbf{x}}_t$ is

$$\Pr(\vec{\mathbf{x}} = \vec{\mathbf{x}}_t) = \frac{1}{\sqrt{(2\pi)^d \, |\mathbf{C}_{xx}|}} \exp\left(-\frac{1}{2}\left(\vec{\mathbf{x}}_t - \vec{\mu}_x\right)^T \mathbf{C}_{xx}^{-1}\left(\vec{\mathbf{x}}_t - \vec{\mu}_x\right)\right)$$

Many more useful facts about multivariable Gaussians may be found online. The Wikepedia article at https://en.wikipedia.org/wiki/Multivariate_normal_distribution is one useful source.

Multiple observations: If we have an initial estimate for the distribution of a multivariable Gaussian random vector $\vec{\mathbf{x}} \sim N(\mu_x^{(0)}, \mathbf{C}_{xx}^{(0)})$ and make N observations $\vec{\mathbf{x}}_k^{(obs)}$ of $\vec{\mathbf{x}}$ where $\vec{\mathbf{x}}_k^{(obs)} \sim N(\vec{\mu}_x^{(k)}, \mathbf{C}_k)$, then our new estimate for the distribution of $\vec{\mathbf{x}}$ is given by

$$\vec{\mathbf{x}}^{(new)} \sim N(\vec{\mu}_X^{(new)}, \mathbf{C}_x^{(new)})$$

$$\mathbf{C}_x^{(new)} = \left(\sum_{0 \le k \le N} \mathbf{C}_k^{-1}\right)^{-1}$$

$$\vec{\mu}_x^{(new)} = \mathbf{C}_x^{(new)}\left(\sum_{0 \le k \le N} \mathbf{C}_k^{-1} \vec{\mu}_x^{(k)}\right)$$

This assignment will make extensive use of the rules for updating conditional multivariable Gaussian distributions. If

$$\vec{\mathbf{x}} = \begin{bmatrix} \vec{\mathbf{x}}_1 \\ \vec{\mathbf{x}}_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \vec{\mu}_1 \\ \vec{\mu}_2 \end{bmatrix}, \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix}\right) \text{ where } \mathbf{C}_{21} = \mathbf{C}_{12}^T$$

then

$$E(\vec{\mathbf{x}}_1 \mid \vec{\mathbf{x}}_2 = \vec{\mathbf{x}}_2^{(obs)}) = \vec{\mu}_1 + \mathbf{C}_{12}\mathbf{C}_{22}^{-1}(\vec{\mathbf{x}}_2^{(obs)} - \vec{\mu}_2)$$
$$\mathrm{cov}(\vec{\mathbf{x}}_1 \mid \vec{\mathbf{x}}_2 = \vec{\mathbf{x}}_2^{(obs)}) = \mathbf{C}_{11} - \mathbf{C}_{12}\mathbf{C}_{22}^{-1}\mathbf{C}_{21}$$

Also, if $\vec{\mathbf{z}} = \vec{\mathbf{x}}_1 + \vec{\mathbf{x}}_2$ then

$$E(\vec{\mathbf{x}}_1 \mid \vec{\mathbf{x}}_1 + \vec{\mathbf{x}}_2 = \vec{\mathbf{z}}) = \vec{\mu}_1 + \left(\mathbf{C}_{11} + \mathbf{C}_{12}\right)\left(\mathbf{C}_{11} + \mathbf{C}_{12} + \mathbf{C}_{21} + \mathbf{C}_{22}\right)^{-1}(\vec{\mathbf{z}} - \vec{\mu}_1 - \vec{\mu}_2)$$
$$\mathrm{cov}(\vec{\mathbf{x}}_1 \mid \vec{\mathbf{x}}_1 + \vec{\mathbf{x}}_2 = \vec{\mathbf{z}}) = \mathbf{C}_{11} - \left(\mathbf{C}_{11} + \mathbf{C}_{12}\right)\left(\mathbf{C}_{11} + \mathbf{C}_{12} + \mathbf{C}_{21} + \mathbf{C}_{22}\right)^{-1}(\mathbf{C}_{11} + \mathbf{C}_{21})$$

Further, if $\mathbf{C}_{12} = \mathbf{0}$ then $\mathrm{cov}(\vec{\mathbf{x}}_1 + \vec{\mathbf{x}}_2 \mid \vec{\mathbf{x}}_1 + \vec{\mathbf{x}}_2 = \vec{\mathbf{z}}) = \mathrm{cov}(\vec{\mathbf{x}}_1) + \mathrm{cov}(\vec{\mathbf{x}}_2)$. Similarly, if $\vec{\mathbf{z}} = \mathbf{A}\vec{\mathbf{x}}_1$ then $E(\vec{\mathbf{z}}) = \mathbf{A}\vec{\mu}_1$ and $\mathrm{cov}(\vec{\mathbf{z}})=\mathbf{A}\mathbf{C}_{11}\mathbf{A}^T$.

Chains of inferences: Given $\vec{\mathbf{a}},\vec{\mathbf{b}},\vec{\mathbf{c}}$ are multivariable Gaussian random vectors with

$$\vec{\mathbf{a}} = \mathbf{M}_{ab}\vec{\mathbf{b}} + \vec{\mathbf{a}}_0,\ \ \mathbf{C}_{a|b} = \mathrm{cov}(\vec{\mathbf{a}} \mid \vec{\mathbf{b}})$$
$$\vec{\mathbf{b}} = \mathbf{M}_{bc}\vec{\mathbf{c}} + \vec{\mathbf{b}}_0,\ \ \mathbf{C}_{b|c} = \mathrm{cov}(\vec{\mathbf{b}} \mid \vec{\mathbf{c}})$$

Then

$$E(\vec{\mathbf{a}} \mid \vec{\mathbf{c}}) = \mathbf{M}_{ab}\mathbf{M}_{bc}E(\vec{\mathbf{c}}) + \mathbf{M}_{ab}\vec{\mathbf{b}}_0 + \vec{\mathbf{a}}_0$$
$$\mathrm{cov}(\vec{\mathbf{a}} \mid \vec{\mathbf{c}}) = \mathbf{C}_{a|b} + \mathbf{M}_{ab}\mathbf{C}_{b|c}\mathbf{M}_{ab}^T$$

Question 1

Suppose that we have linear model $\mathbf{B}_k\vec{\beta}^{(k)} + \mathbf{D}_k\vec{\delta}_k = \vec{\phi}_k$ and where $\mathbf{B}_k$ is an $m \times n_\beta$ matrix, $\mathbf{D}_k$ is an $m \times n_\delta$ matrix, $\vec{\beta}^{(k)}$ is an n_β-vector, and $\vec{\delta}_k$ is a random n_δ-vector with

$$\vec{\beta}^{(k)} \sim N(\vec{\mu}_\beta^{(k)}, \mathbf{C}_{\beta\beta}^{(k)}) \quad \vec{\delta}_k \sim N(\vec{\mathbf{0}}, \mathbf{C}_{\delta\delta,k}) \quad Cov(\vec{\beta}^{(k)}, \delta_k) = \mathbf{0}$$

$$\vec{\phi}_k \sim N(\mathbf{B}_k\vec{\mu}_\beta^{(k)}, \mathbf{C}_{\phi\phi}^{(k)}) \quad Cov(\vec{\phi}_j, \vec{\phi}_k) = \mathbf{0} \text{ for } j \neq k$$

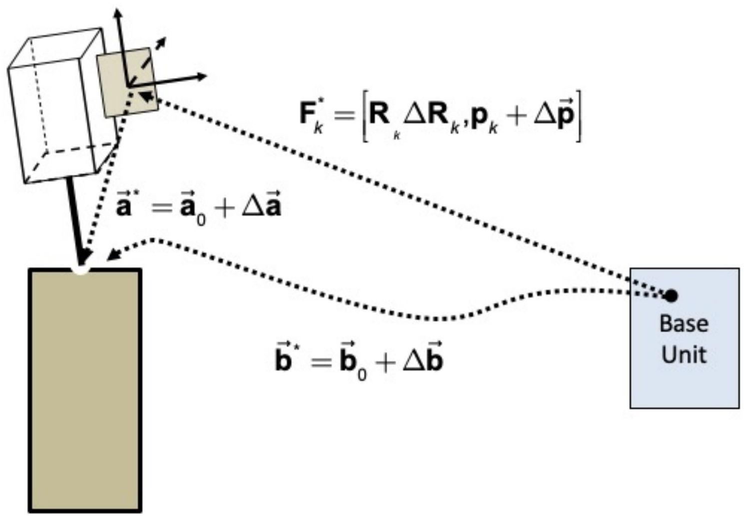
Suppose also that we have initial distributions $\vec{\beta}^{(0)} \sim N(\vec{\mu}_\beta^{(0)}, \mathbf{C}_{\beta\beta}^{(0)})$ $\vec{\delta}_k \sim N(\vec{\mathbf{0}}, \mathbf{C}_{\delta\delta}^{(0)})$ and a set of N_{obs} observed values $\left\{\cdots, (\mathbf{B}_k, \mathbf{D}_k, \vec{\phi}_k^{(obs)}), \cdots\right\}$. Our goal is to use these observations to find an updated distribution $\vec{\beta}^{(N_{obs})} \sim N(\mu_\beta^{(N_{obs})}, \mathbf{C}_{\beta\beta}^{(N_{obs})})$.

A. (5 points) Given an observation $(\mathbf{B}_k, \mathbf{D}_k, \vec{\phi}_k^{(obs)})$, what are the mean and covariance of $\vec{\beta}^{(k+1)} \sim N(\vec{\mu}_\beta^{(k+1)}, \mathbf{C}_{\beta\beta}^{(k+1)})$ where $\vec{\beta}^{(k+1)} = (\vec{\beta}^{(k)} \mid \mathbf{B}_k, \mathbf{D}_k, \vec{\phi}_k^{(obs)})$? I.e., give expressions for $\vec{\mu}_\beta^{(k+1)}$ and $\mathbf{C}_{\beta\beta}^{(k+1)}$ in terms of $\mathbf{B}_k, \mathbf{D}_k, \vec{\phi}_k^{(obs)}, \mu_\beta^{(k)}$, and $\mathbf{C}_{\beta\beta}^{(k)}$.

B. (5 points) Describe a sequential algorithm for estimating the distribution of random variable $\beta^{(N)} \sim N(\vec{\mu}_\beta^{(N)}, \mathbf{C}_{\beta\beta}^{(N)})$ given the assumptions above by incrementally updates as each observation becomes available.

Question 2

Pointing device Re-calibration



Consider the problem of recalibrating a previously calibrated pointing device. In this scenario, the system knows approximate values $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ for the tip offset and pivot dimple positions, respectively. These values can be considered to be random variables, with

$$\vec{\mathbf{a}} \sim N(\vec{\mathbf{a}}_0, \mathbf{C}_{aa}^{(0)})$$
$$\vec{\mathbf{b}} \sim N(\vec{\mathbf{b}}_0, \mathbf{C}_{bb}^{(0)})$$
$$\mathrm{cov}(\vec{\mathbf{a}}, \vec{\mathbf{b}}) = \mathbf{C}_{ab}^{(0)}$$

where $\vec{\mathbf{a}}_0$ and $\vec{\mathbf{b}}_0$ are the computer's current values for $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$. If we consider $\Delta\vec{\mathbf{a}}^{(0)}$ and $\Delta\vec{\mathbf{b}}^{(0)}$ to be random variables representing the error in the computer's current values, we also have

$$\Delta\vec{\mathbf{a}} \sim N(\vec{\mathbf{0}}, \mathbf{C}_{aa}^{(0)}) \quad \Delta\vec{\mathbf{b}} \sim N(\vec{\mathbf{0}}, \mathbf{C}_{bb}^{(0)}) \quad \mathrm{Cov}(\Delta\vec{\mathbf{a}}, \Delta\vec{\mathbf{b}}) = \mathrm{cov}(\Delta\vec{\mathbf{b}}, \Delta\vec{\mathbf{a}}) = \mathbf{C}_{ab}^{(0)}$$

A collection of poses $\mathbf{F}_k$ are available with the pointer tip in the dimple. There is some noise in the $\mathbf{F}_k$ values, so that the actual values are

$$\mathbf{F}_k^* = \mathbf{F}_k \Delta\mathbf{F}_k = \mathbf{F}_k \bullet \left[\Delta\mathbf{R}_k, \Delta\vec{\mathbf{p}}_k\right] \approx \left[\mathbf{I} + sk(\vec{\alpha}_k), \vec{\varepsilon}_k\right].$$

Further, we have the following information about $\vec{\alpha}_k$ and $\vec{\varepsilon}_k$,

$$\vec{\eta}_k = \begin{bmatrix} \vec{\alpha}_k \\ \vec{\varepsilon}_k \end{bmatrix} \sim N(\vec{0}_6, \mathbf{C}_{\eta\eta}^{(k)}) \text{ and } \mathbf{C}_{\eta\eta}^{(k)} = \begin{bmatrix} \mathbf{C}_{\alpha\alpha}^{(k)} & \mathbf{C}_{\alpha\varepsilon}^{(k)} \\ \mathbf{C}_{\varepsilon\alpha}^{(k)} & \mathbf{C}_{\varepsilon\varepsilon}^{(k)} \end{bmatrix}$$

For the purposes of this problem, we will assume that the random variables $\vec{\eta}_k$ are independent and identically distributed. I.e., $\text{cov}(\vec{\eta}_j, \vec{\eta}_k) = \mathbf{0}_{6x6}$ for $i \neq j$. Also, note that the $\vec{\eta}_k$ are independent of $\Delta\vec{\mathbf{a}}$ and $\Delta\vec{\mathbf{b}}$, so that $\text{cov}(\vec{\eta}_k, \Delta\vec{\mathbf{a}}) = \text{cov}(\vec{\eta}_k, \Delta\vec{\mathbf{b}}) = \mathbf{0}$.

Our goal is to use this information to update our values for $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$, together with our estimates of the covariances.

A. (5 points) Write an equation for $\mathbf{F}_k^* \vec{\mathbf{a}}^* - \mathbf{b}^*$ in terms of $\mathbf{R}_k, \vec{\mathbf{p}}_k, \vec{\mathbf{d}}, \Delta\mathbf{R}_k, \Delta\vec{\mathbf{p}}_k, \Delta\vec{\mathbf{a}}, \Delta\vec{\mathbf{b}}$.

B. (5 points) Rewrite your equation by using the approximation $\Delta\mathbf{F}_k \approx \left[\mathbf{I} + sk(\vec{\alpha}_k), \vec{\varepsilon}_k\right]$ and eliminating 2nd order terms.

C. (10 points) Rearrange your answer into a matrix form suitable for applying the methods of Question 1 to produce estimates of $\Delta\vec{\mathbf{a}}$ and $\Delta\vec{\mathbf{b}}$ that can be used to update our values for $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$, as well as the corresponding co-variances. This will be done in two stages. For Question 2.B, you should define a random vector

$$\Delta\vec{\mathbf{d}} = \begin{bmatrix} \Delta\vec{\mathbf{a}} \\ \Delta\vec{\mathbf{b}} \end{bmatrix} \text{ with } \mathbf{C}_{dd}^{(0)} = \text{cov}(\Delta\vec{\mathbf{d}}, \Delta\vec{\mathbf{d}}) = \begin{bmatrix} \mathbf{C}_{aa}^{(0)} & \mathbf{C}_{ab}^{(0)} \\ \mathbf{C}_{ab}^{(0)} & \mathbf{C}_{bb}^{(0)} \end{bmatrix}$$

and produce expressions for producing a new estimate $\Delta\vec{\mathbf{d}}^{(new)}$, where $\Delta\vec{\mathbf{d}}^{(new)} \sim N\left(E(\Delta\vec{\mathbf{d}}^{(new)}), \text{cov}(\Delta\vec{\mathbf{d}}^{(new)}, \Delta\vec{\mathbf{d}}^{(new)})\right)$. I.e., produce expressions for

$$E(\Delta\vec{\mathbf{d}}^{(new)}) = E(\Delta\vec{\mathbf{d}} \mid \mathbf{R}_k, \vec{\mathbf{p}}_k, \vec{\mathbf{a}} = \vec{\mathbf{a}}_0, \vec{\mathbf{b}} = \vec{\mathbf{b}}_0)$$
$$\text{cov}(\Delta\vec{\mathbf{d}}^{(new)}) = \text{cov}(\Delta\vec{\mathbf{d}} \mid \mathbf{R}_k, \vec{\mathbf{p}}_k, \vec{\mathbf{a}} = \vec{\mathbf{a}}_0, \vec{\mathbf{b}} = \vec{\mathbf{b}}_0)$$

As part of your answer, you should define matrices $\mathbf{A}_k$ and $\mathbf{D}_k$ in terms of $\{\mathbf{R}_k, \vec{\mathbf{p}}_k, \vec{\mathbf{a}}_0, \vec{\mathbf{b}}_0\}$ and a random vector $\vec{\mathbf{e}}_k$ and

$$\begin{aligned} \vec{\mathbf{e}}_k &= \left(\mathbf{F}_k^* \vec{\mathbf{a}}^* - \vec{\mathbf{b}}^*\right) - \left(\mathbf{F}_k \vec{\mathbf{a}} - \vec{\mathbf{b}}\right) \\ &= \mathbf{A}_k \vec{\eta}_k + \mathbf{D}_k \Delta\vec{\mathbf{d}} \end{aligned}$$

For a given $\mathbf{R}_k, \vec{\mathbf{p}}_k, \vec{\mathbf{a}} = \vec{\mathbf{a}}_0, \vec{\mathbf{b}} = \vec{\mathbf{b}}_0$, we have $\vec{\mathbf{e}}_k^{(obs)} = \mathbf{F}_k \vec{\mathbf{a}}_0 - \vec{\mathbf{b}}_0 = \mathbf{R}_k \vec{\mathbf{a}}_0 + \vec{\mathbf{p}}_k - \vec{\mathbf{b}}_0$

As part of your answer, you will also need to develop expressions $\mathbf{C}_{ee}^{(k)} = \text{cov}(\vec{\mathbf{e}}_k, \vec{\mathbf{e}}_k)$ in terms of $\mathbf{A}_k$ and $\mathbf{D}_k$ and the other covariances in the problem. You can then use $\mathbf{C}_{ee}^{(k)}$ and $\mathbf{C}_{dd} = \text{cov}(\Delta\vec{\mathbf{d}}, \Delta\vec{\mathbf{d}})$ in your answer, which may also include $\mathbf{A}_k$ and/or $\mathbf{D}_k$. You will expand things further in the next question.

D. (10 points) Now expand your answer to Question 2.C to provide expressions for

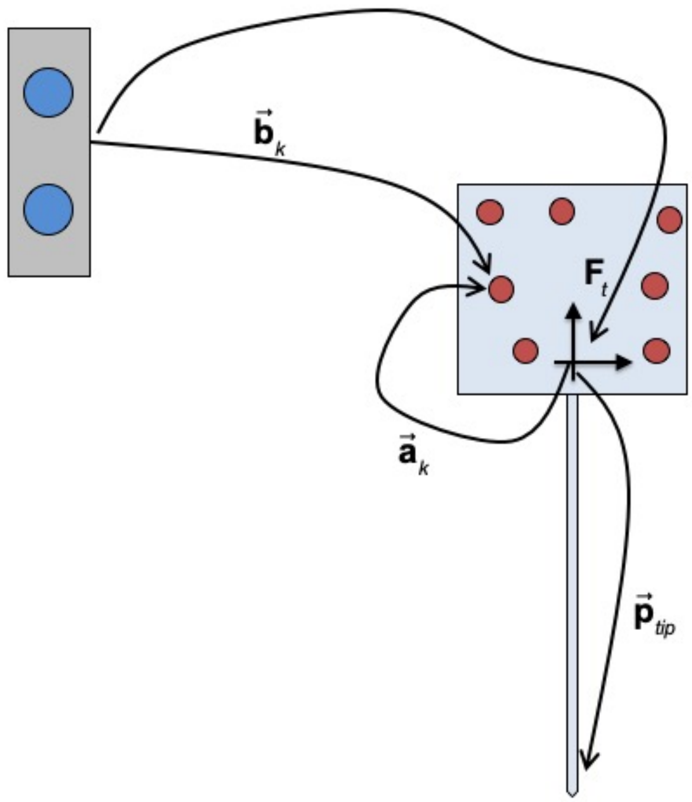
$$E(\Delta\vec{\mathbf{a}}^{(new)}) = E\left(\Delta\vec{\mathbf{a}} \mid \{\mathbf{R}_k, \vec{\mathbf{p}}_k, \vec{\mathbf{e}}_k^{(obs)}\}\right)$$
$$E(\Delta\vec{\mathbf{b}}^{(new)}) = E\left(\Delta\vec{\mathbf{b}} \mid \{\mathbf{R}_k, \vec{\mathbf{p}}_k, \vec{\mathbf{e}}_k^{(obs)}\}\right)$$

$$\mathrm{cov}(\Delta\vec{\mathbf{a}})^{(new)} = \mathrm{cov}(\Delta\vec{\mathbf{a}} \mid \{\mathbf{R}_k, \vec{\mathbf{p}}_k, \vec{\mathbf{a}}_0, \vec{\mathbf{b}}_0\})$$
$$\mathrm{cov}(\Delta\vec{\mathbf{b}})^{(new)} \leftarrow \mathrm{cov}(\Delta\vec{\mathbf{b}} \mid \{\mathbf{R}_k, \vec{\mathbf{p}}_k, \vec{\mathbf{a}}_0, \vec{\mathbf{b}}_0\})$$
$$\mathrm{cov}(\Delta\vec{\mathbf{a}}, \Delta\vec{\mathbf{b}})^{(new)} \leftarrow \mathrm{cov}\left((\Delta\vec{\mathbf{a}} \mid \{\mathbf{R}_k, \vec{\mathbf{p}}_k, \vec{\mathbf{a}}_0, \vec{\mathbf{b}}_0\}),\ (\Delta\vec{\mathbf{b}} \mid \{\mathbf{R}_k, \vec{\mathbf{p}}_k, \vec{\mathbf{a}}_0, \vec{\mathbf{b}}_0\})\right)$$

In your answers, you should provide an expression for $\mathbf{C}_{ee}^{(k)} = \mathrm{cov}(\vec{\mathbf{e}}_k, \vec{\mathbf{e}}_k)$ in terms of $\mathbf{R}_k, \vec{\mathbf{p}}_k, \vec{\mathbf{a}}_0, \vec{\mathbf{b}}_0, \mathbf{C}_{\alpha\alpha}, \mathbf{C}_{\alpha\varepsilon}, \mathbf{C}_{\varepsilon\varepsilon}, \mathbf{C}_{aa}, \mathbf{C}_{ab}, \mathbf{C}_{bb}$. You can then use $\mathbf{C}_{ee}^{(k)}$ in your answer expressions.

E. (5 points) Suppose that the pivoting operation is imperfect so that $\mathbf{F}_k^* \vec{\mathbf{a}}^* - \vec{\mathbf{b}}^* = \Delta\vec{\mathbf{c}}_k$ where $\Delta\vec{\mathbf{c}}_k \sim N(\vec{\mathbf{0}}, \mathbf{C}_{cc})$ and $\mathrm{cov}(\Delta\vec{\mathbf{c}}_k, \Delta\vec{\mathbf{d}}) = \mathrm{cov}(\vec{\eta}_k, \Delta\mathbf{d}) = \mathrm{cov}(\vec{\eta}_k, \Delta\vec{\mathbf{c}}_k) = \mathbf{0}$. How will this affect your answer to Question 2.0? Here, I am not looking for you to expand out the whole answer again, but just explain the effect on $\vec{\mathbf{e}}_k$ and $\mathbf{C}_{ee}^{(k)}$. **Note:** essentially, we are asking for you to explain what will happen to your formulas for the mean and covariance of $\vec{\mathbf{e}}_k$.

Question 3



Now consider the situation above. We have created a pointer object by sticking a number of markers onto the proximal part of the pointer at positions $\vec{\mathbf{a}}_k$ with respect to the pointer coordinate system. The position of the pointer tip relative to the pointer coordinate system is $\vec{\mathbf{p}}_{tip}$. The true position of each marker $\vec{\mathbf{a}}_k$ relative to the a tracking device is $\vec{\mathbf{b}}_k^* = \vec{\mathbf{b}}_k + \Delta\vec{\mathbf{b}}_k$ where $\Delta\vec{\mathbf{b}}_k \sim N(\vec{\mathbf{0}}, \mathbf{C}_{bb})$. At time t the pose $\mathbf{F} = [\mathbf{R}, \mathbf{p}]$ has been computed from the $\{\cdots(\vec{\mathbf{a}}_k, \vec{\mathbf{b}}_k)\cdots\}$ values. The true pose will thus have the relationship $\mathbf{F}^*\vec{\mathbf{a}}_k = \vec{\mathbf{b}}_k^*$. Where $\mathbf{F}_t^* = \mathbf{F}_t\Delta\mathbf{F} = [\mathbf{R}_t\Delta\mathbf{R}_t, \vec{\mathbf{p}}_t + \mathbf{R}_t\Delta\vec{\mathbf{p}}_k]$ with $\Delta\mathbf{R}_t \approx \mathbf{I} + sk(\vec{\alpha}_t)$ $\Delta\vec{\mathbf{p}}_t = \vec{\varepsilon}_t$, with

$$\vec{\eta}_t = \begin{bmatrix} \vec{\alpha}_t \\ \vec{\varepsilon}_t \end{bmatrix} \sim N\left(\begin{bmatrix} \vec{\mathbf{0}} \\ \vec{\mathbf{0}} \end{bmatrix}, \mathbf{C}_{\eta\eta} = \begin{bmatrix} \mathbf{C}_{\alpha\alpha} & \mathbf{C}_{\alpha\varepsilon} \\ \mathbf{C}_{\varepsilon\alpha} & \mathbf{C}_{\varepsilon\varepsilon} \end{bmatrix} \right)$$

Note: Several of these questions assume a knowledge of the properties of determinants, which you should have learned in your linear algebra course. If you need a refresher, then there is an excellent Wikipedia article at https://en.wikipedia.org/wiki/Determinant. A few useful facts for real positive definite $d \times d$ matrices $\mathbf{A}$ and $\mathbf{B}$ are

$$
\begin{aligned}
&|\mathbf{A}^T| = |\mathbf{A}| \\
&|\mathbf{A}^{-1}| = |\mathbf{A}|^{-1} \\
&|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}| \quad \Rightarrow \quad |\mathbf{I}| = 1 \\
&|c\mathbf{A}| = c^d\,|\mathbf{A}| \\
&|\mathbf{A}+\mathbf{B}| \ge |\mathbf{A}| + |\mathbf{B}|
\end{aligned}
$$

A. (5 points) Give an expression for $\vec{\mathbf{e}}_t \approx \mathbf{F}_t^* \vec{\mathbf{p}}_{tip} - \mathbf{F}_t \vec{\mathbf{p}}_{tip}$ in terms of $\vec{\alpha}_t$ and $\vec{\varepsilon}_t$.

B. (10 points) Now consider that $\vec{\mathbf{e}}_t$ is a random vector with $\vec{\mathbf{e}}_t \sim N(\vec{\mu}_t, \mathbf{C}_{ee})$. Provide expressions for $E(\vec{\mathbf{e}}_t \mid \vec{\eta}_t)$ and $\mathrm{cov}(\vec{\mathbf{e}}_t \mid \vec{\eta}_t)$ in terms of $\mathbf{R}_t, \vec{\mathbf{p}}_t, \vec{\mathbf{p}}_{tip}$, and the elements of $\mathbf{C}_{\eta\eta}$.

C. (5 points) Suppose that we know that $\left|\mathbf{C}_{\eta\eta}\right| \le \sigma_\eta$, what can we say about $\mathbf{C}_{ee}$?

D. (5 points) Give expressions for $\vec{\mu}_t = E(\Delta\vec{\mathbf{b}}_k \mid \vec{\eta}_t)$ and $\mathbf{C}_k = \mathrm{cov}(\Delta\vec{\mathbf{b}}_k \mid \vec{\eta}_t)$ in terms of $\mathbf{R}_t, \vec{\mathbf{p}}_t, \vec{\mathbf{a}}_k, \vec{\mathbf{b}}_k$ and the elements of $\mathbf{C}_{\eta\eta}$

E. (5 points) Now give expressions for $E\left(\vec{\eta}_t \mid \{\cdots,(\mathbf{R}_t,\vec{\mathbf{p}}_t,\vec{\mathbf{a}}_k,\vec{\mathbf{b}}_k),\cdots\}\right)$ and $\mathrm{cov}\left(\vec{\eta}_t \mid \{\cdots,(\mathbf{R}_t,\vec{\mathbf{p}}_t,\vec{\mathbf{a}}_k,\vec{\mathbf{b}}_k),\cdots\}\right)$.

F. (5 points) What can we say about $|\mathrm{cov}\left(\vec{\eta}_t \mid \{\cdots,(\mathbf{R}_t,\vec{\mathbf{p}}_t,\vec{\mathbf{a}}_k,\vec{\mathbf{b}}_k),\cdots\}\right)|$? **Note:** the previous question asked for an estimate of the distribution of $\vec{\eta}_t$ based on a set of observations for the $\vec{\mathbf{b}}_k$ at time t. Here we are asking for the determinant of the covariance. This is an important value for estimating the probability of any particular pose at time t.

G. (5 points) Now, what can we say about $|\mathrm{cov}\left(\vec{\mathbf{e}}_t \mid \{\cdots,(\mathbf{R}_t,\vec{\mathbf{p}}_t,\vec{\mathbf{b}}_k),\cdots\}\right)|$? **Hint:** Use the rule for chains of inferences for multivariable Gaussians.

H. (5 points) Now suppose that there is some uncertainty in the positions of the $\vec{\mathbf{a}}_k$ relative to the pointer coordinate system, so that $\vec{\mathbf{a}}_k^* = \vec{\mathbf{a}}_k + \Delta\vec{\mathbf{a}}_k$. However, we know that $|\mathbf{C}_{bb}| \le \sigma_b$. Outline a method to produce an improved estimate for the $\vec{\mathbf{a}}_k$ values using the apparatus available. I.e., you should make some measurements and produce an updated distribution $\vec{\mathbf{a}}_k^{(new)} \sim N\left(\vec{\mu}_k^{(new)}, \mathbf{C}_k^{(new)}\right)$ with

$$
\begin{aligned}
\vec{\mu}_k^{(new)} &= E(\vec{\mathbf{a}}_k^{(new)}) \leftarrow \vec{\mathbf{a}}_k + E(\Delta\vec{\mathbf{a}}_k \mid \text{some data}) \\
\mathbf{C}_k^{(new)} &= \mathrm{cov}(\vec{\mathbf{a}}_k^{(new)}) \leftarrow E\left((\vec{\mathbf{a}}_k^{(new)} - \vec{\mu}_k^{(new)})(\vec{\mathbf{a}}_k^{(new)} - \vec{\mu}_k^{(new)})^T\right)
\end{aligned}
$$

I. (5 points) What can you say about $|\mathbf{C}_k^{(new)}|$?

Question 4

Suppose that we have two measurements of a function $\vec{\mathbf{f}}(\vec{\mathbf{x}})$, with $\vec{\mathbf{f}}(\vec{\mathbf{x}}_1) = \vec{\mathbf{f}}_1$ and $\vec{\mathbf{f}}(\vec{\mathbf{x}}_2) = \vec{\mathbf{f}}_2$. We wish to estimate $\vec{\mathbf{f}}(\vec{\mathbf{x}})$ by linear interpolation over the interval $[\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2]$:

$$\vec{\mathbf{f}}^{(e)}(\vec{\mathbf{x}}) = \lambda \mathbf{f}_1 + \vec{\mu}\vec{\mathbf{f}}_2$$
$$\vec{\mathbf{x}} = \lambda \vec{\mathbf{x}}_1 + \vec{\mu}\vec{\mathbf{x}}_2$$
$$\lambda + \mu = 1$$

A. (10 points) Suppose also that we know that $\left\|\nabla \vec{\mathbf{f}}\right\| \le \phi$. What is the maximum value of $\varepsilon(\vec{\mathbf{x}}) = \left\|\vec{\mathbf{e}}(\vec{\mathbf{x}}) = \vec{\mathbf{f}}(\vec{\mathbf{x}}) - \vec{\mathbf{f}}^{(e)}(\vec{\mathbf{x}})\right\|$ over the interval $[\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2]$? **Hint:** There will be some values $0 \le \lambda \le 1$ and $\mu = 1 - \lambda$ such that $\varepsilon(\lambda \vec{\mathbf{x}}_1 + \mu \vec{\mathbf{x}}_2)$ is maximized.

B. (5 points) Suppose now that there is some error in our two measurements, so that

$$\varepsilon_1 = \left\|\vec{\mathbf{e}}_1 = \vec{\mathbf{f}}^*(\vec{\mathbf{x}}_1) - \vec{\mathbf{f}}(\vec{\mathbf{x}}_1)\right\|$$
$$\varepsilon_2 = \left\|\vec{\mathbf{e}}_2 = \vec{\mathbf{f}}^*(\vec{\mathbf{x}}_2) - \vec{\mathbf{f}}(\vec{\mathbf{x}}_2)\right\|$$

What is the maximum value for $\varepsilon(\vec{\mathbf{x}})$ over the interval?

C. ((5 points) What is the maximum distance $\left\|\vec{\mathbf{x}}_1 - \vec{\mathbf{x}}_2\right\|$ such that $\left\|\varepsilon(\vec{\mathbf{x}})\right\| \le \delta$ over the interval for some value δ?

D. (5 points) Suppose instead that the errors are random, with $\vec{\mathbf{e}}(\vec{\mathbf{x}}) \sim N\left(\vec{\mathbf{0}}, \mathbf{C}_e(\vec{\mathbf{x}})\right)$ with

$$\mathbf{C}_e(\vec{\mathbf{x}}) = diag\left(\left[\sigma(\vec{\mathbf{x}}), \cdots, \sigma(\vec{\mathbf{x}})\right]\right) = \begin{bmatrix} \sigma(\vec{\mathbf{x}}) & & & & \\ & \ddots & & & \\ & & \sigma(\vec{\mathbf{x}}) & & \\ & & & \ddots & \\ & & & & \sigma(\vec{\mathbf{x}}) \end{bmatrix}.$$

Suppose that we know $\mathbf{C}_e(\vec{\mathbf{x}}_1) = \mathbf{C}_1$ and $\mathbf{C}_e(\vec{\mathbf{x}}_2) = \mathbf{C}_2$. Give an expression for $\mathbf{C}_e(\lambda \vec{\mathbf{x}}_1 + (1-\lambda)\vec{\mathbf{x}}_2)$.

E. (5 points) What is the maximum distance $\left\|\vec{\mathbf{x}}_1 - \vec{\mathbf{x}}_2\right\|$ such that $\left|\mathbf{C}_e(\vec{\mathbf{x}})\right| \le \delta$ over the interval for some value δ?