Solved Exercises in Physical Measurements

Notes from the Exercises component of the third-year undergraduate course *Fizikalna merjenja* (Physical Measurements), led by doc. dr. Gregor Kladnik at the Faculty of Mathematics and Physics at the University of Ljubljana in the academic year 2020-21. The course covers a selection of practical topics related to physical measurement. Credit for the material covered in these notes is due to professor Kladnik, while the voice, typesetting, and translation to English in this document are my own.

Disclaimer: Mistakes—both simple typos and legitimate errors—are likely. Keep in mind that these are the notes of an undergraduate student in the process of learning the material himself—take what you read with a grain of salt. If you find mistakes and feel like telling me, for example by email, I'll be happy to hear from you, even for the most trivial of errors.

Elijan J. Mastnak
Faculty of Mathematics and Physics, University of Ljubljana
Summer semester 2020-2021

Contents

1 First Exercise Set

1.1 Propagation of Uncertainty

- Consider a quantity u that is a function of n known input quantities $x_1, x_2, \ldots, x_n$,

$$u = f(x_1, x_2, \ldots, x_n).$$

Our goal, given the uncertainties $\delta x_1, \delta x_2, \ldots, \delta x_n$ in the input quantities, is to find the corresponding uncertainty δu in the derived quantity u.

- In general terms the situation may be written

$$u + \delta u = f(x_1 + \delta x_1, \ldots, x_n + \delta x_n) \tag{1.1}$$

Before going further, we introduce the more compact vector notation

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad \delta\boldsymbol{x} = \begin{pmatrix} \delta x_1 \\ \vdots \\ \delta x_n \end{pmatrix}, \qquad \left(\frac{\partial f}{\partial \boldsymbol{x}}\right) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}_{\boldsymbol{x}}.$$

In vector notation, Equation 1.1 then reads

$$u + \delta u = f(\boldsymbol{x} + \delta\boldsymbol{x}).$$

- Assuming $\delta x_i \ll x_i$, we then Taylor expand $u + \delta u$ to first order about $u = f(\boldsymbol{x})$, which reads

$$u + \delta u = f(\boldsymbol{x}) + \left(\frac{\partial f}{\partial x_i}\right)_{\boldsymbol{x}} \delta x_1 + \cdots + \left(\frac{\partial f}{\partial x_n}\right)_{\boldsymbol{x}} \delta x_n + \mathcal{O}(\delta x_i),$$

where the partial deivatives are evaulated at $\boldsymbol{x}$ **TODO:** maybe switch to $\boldsymbol{x}_0$. The terms u and $f(\boldsymbol{x}) = u$ cancel from both sides of the equation, leaving

$$\begin{aligned} \delta u &= \left(\frac{\partial f}{\partial x_i}\right)_{\boldsymbol{x}} \delta x_1 + \cdots + \left(\frac{\partial f}{\partial x_n}\right)_{\boldsymbol{x}} \delta x_n + \mathcal{O}(\delta x_i) \\ &= \delta\boldsymbol{x}^\top \cdot \left(\frac{\partial f}{\partial \boldsymbol{x}}\right)_{\boldsymbol{x}}. \end{aligned}$$

The corresponding relative uncertainty in δu is

$$\frac{\Delta u}{u} = \frac{1}{f(\boldsymbol{x})} \Delta\boldsymbol{x}^\top \left(\frac{\partial f}{\partial \boldsymbol{x}}\right)_{\boldsymbol{x}}. \tag{1.2}$$

- *Important!* So far we have assumed that all δ_i are positive!

We have done: we increase value of x_i by δx_i and ask how δu increases as a result.

This is not a general treament for $x_i \pm x_i$.

1.2 The Fundamental Frequency of a Harmonic String

A harmonic string of length l and mass m is fixed between two points with a tension force of magnitude F. Estimate the relative precision $(\delta\nu_0)/\nu_0$ of the fundamental frequency.

- We will denote the string's amplitude during oscillation by y.
- When the string's amplitude changes from its equilibrium postion, the string stretches slightly, and its length changes from l to $l + \delta l$, while the tension force increases from F to $F + \delta F$.
- Our first step is to find a functional expression for the oscillation frequency ν. For simple harmonic motion, we assume ν is not a function of x and use the general wave formula
$$c = \nu\lambda \implies \nu = \frac{c}{\lambda},$$
where λ is the wave's wavelength. For the fundamental frequency we have $\lambda_0 = 2l$ (Fizika 1). Again quoting from introductory wave mechanics, the speed of standing waves (TF?) on a string with linear mass density μ is
$$c = \sqrt{\frac{F}{\mu}}$$
where $\mu = m/l$ is the string's linear mass density, calculated in the *equilibrium* position. (Note that μ changes as the string oscillates and it's length l changes.)
- The expression for the fundamental wave frequency ν_0 is then
$$\nu_0 = \frac{c}{\lambda_0} = \frac{1}{2l} \cdot \sqrt{\frac{F}{\mu}} \equiv f(l, F, \mu). \tag{1.3}$$
For compactness, we will write our input quantities as the vector
$$\boldsymbol{x} = \begin{pmatrix} l \\ f \\ \mu \end{pmatrix}$$
- Our goal is to find $\frac{\delta\nu_0}{\nu_0}$. From Equation 1.2, the general expression is
$$\frac{\delta\nu}{\nu} = \frac{1}{f(\boldsymbol{x})} \delta\boldsymbol{x}^\top \left(\frac{\partial f}{\partial \boldsymbol{x}}\right). \tag{1.4}$$
Goal: find the effect of oscillation amplitude y on $(\delta\nu)/\nu$ indirectly, by first finding the effect of y on length l and force F.
- Recall that the string length l arises from the relationship $\lambda_0 = 2l$. But (for harmonic oscillation) wavelength λ does not depend on amplitude y.
So we could really write $\boldsymbol{x} = \big(F, \mu\big)^\top$ and $\nu_0 = f(F, \mu)$.

- We define the vectors...

$$\delta \boldsymbol{x} = \begin{pmatrix} \delta F \\ \delta \mu \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} \frac{\partial f}{\partial \boldsymbol{x}} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial F} \\ \frac{\partial f}{\partial \mu} \end{pmatrix}$$

  Using the expression for ν_0 in Equation 1.3, we then compute

$$\frac{\partial f}{\partial F} = \frac{1}{4l}\frac{1}{\sqrt{\mu F}} \qquad \text{and} \qquad \frac{\partial f}{\partial \mu} = -\frac{1}{4l}\frac{\sqrt{F}}{\mu^{3/2}}.$$

- From Equation 1.4, the relative uncertainty in fundamental frequency is then

$$\begin{aligned} \frac{\delta \nu}{\nu} &= \frac{1}{f(\boldsymbol{x})} \delta \boldsymbol{x}^\top \left(\frac{\partial f}{\partial \boldsymbol{x}} \right) = \left(2l\sqrt{\frac{\mu}{F}} \right) \cdot (\delta F, \delta \mu)^\top \cdot \frac{1}{4l} \begin{pmatrix} \frac{1}{\sqrt{\mu F}} \\ -\frac{\sqrt{F}}{\mu^{3/2}} \end{pmatrix} \\ &= \frac{1}{2}\left(\frac{\delta F}{F} - \frac{\delta \mu}{\mu} \right). \end{aligned}$$

- Next is to write changes in F and μ in terms of the amplitude y.
  - First, using $\mu = m/l$, where the string mass m is constant, the relative change in mass density corresponding to relative change in length $(\delta l)/l$ is simply

$$\frac{\delta \mu}{\mu} = -\frac{\delta l}{l}.$$

  - We find the relative change in force $\frac{\delta F}{F}$ from the general elastomechanical relationship

$$\frac{\delta F}{S} = E\frac{\delta l}{l} \implies \frac{\delta F}{F} = \frac{ES}{F}\frac{\delta l}{l},$$

    where E and S are the string's Young's modulus and cross-sectional area.

- Using the just-derived results, the relative uncertainty in frequency is then

$$\frac{\delta \nu}{\nu} = \frac{1}{2}\left(\frac{\delta F}{F} - \frac{\delta \mu}{\mu} \right) = \frac{1}{2}\left(\frac{ES}{F}\frac{\delta l}{l} + \frac{\delta l}{l} \right) = \frac{1}{2}\frac{\delta l}{l}\left(1 + \frac{ES}{F} \right)$$

  This is almost the desired result—it tells us how relative uncertainty in ν depends on the string length l, together with the constant parameters E, S and F and l. We now just want to write this in terms of amplitude y.

Relating Length l and Amplitude y

- We assume the oscillation amplitude is small, i.e. that $y \ll l$. Note that this is usually realistic: consider a guitar string where amplitude is much less than the string length.

- Length after displacement is $l+\delta l$. The length of half of the string is thus $(l+\delta l)/2$

  Model: (to avoid elliptic integrals) We model the dispaced string as an isosceles triangle with base l, equal side lengths $(l+\delta l)/2$, and height y.

The Pythagorean theorem applied to one half of the triangle gives

$$\left(\frac{l+\delta l}{2}\right)^2 \approx y^2 + \left(\frac{l}{2}\right)^2$$

We then multiply through by four and rearrange to get

$$l^2 + 2\delta l \cdot l + (\delta l)^2 = 4y^2 + l^2$$

Assuming $\delta l \ll l$, we neglect the second-order term $(\delta l)^2$.[1] We then divide through by l^2 to get the desired relationship between δl and y:

$$\frac{\delta l}{l} = 2\left(\frac{y}{l}\right)^2.$$

- The relative uncertainty in fundamental frequency is then

$$\frac{\delta \nu}{\nu} = \frac{1}{2}\frac{\delta l}{l}\left(1 + \frac{ES}{F}\right) = \left(\frac{y}{l}\right)^2\left(1 + \frac{ES}{F}\right),$$

which is written only in terms of the displacement y and the constant parameters F and initial length l, and in terms of amplitude y.

[1] This step is not really thorough, because δl is of the same order as oscillation amplitude y, and we aren't neglecting y. That's okay though, this is just an approximation.

2 Second Exercise Set

2.1 Example: Measuring a Musical Semitones

How long to listen to a note with $\nu_0 = 440\,\mathrm{Hz}$ to be able to determine the tone to a half-tone accuracy.

- We will use an uncertainty principle of the form
$$\Delta\nu\Delta t \geq \frac{1}{2} \gtrsim 1$$
Time refers to the length of time we observe the signal. When used only as an estimate, we replace $1/2 \approx 1$.

- Our goal is to find time Δt for the time of the signal, such that $\delta\nu$ is accurate to a half-tone. This is simply, using the rough uncertainty principle estimate.
$$\Delta t \gtrsim \frac{1}{\Delta\nu}$$

- The problem reduces to finding the correct expression for $\Delta\nu$. Note that ν_0 in practice would not be known exactly—this problem is artificial.
The half-tone has a well-defined quantitative value, using the 12 tone equal temperament tuning system. The ratio between any two neighboring tones is constant.
$$\frac{\widehat{\nu}_{i+1}}{\widehat{\nu}} = \sqrt[12]{2}$$
Hats just denote well-defined semitones.

- The next tone above ν_0 is ν_1. Our expression for $\Delta\nu$ is thus
$$\Delta\nu = \widehat{\nu}_1 - \widehat{\nu}_0 = x\widehat{\nu}_0 - \widehat{\nu}_0 = \widehat{\nu}_0(x-1) = \nu_0(x-1)$$
Note the distinction between generic semitones and the well-defined ν_0.
In 12-TET, the ratio between successive tones is
$$x = \sqrt[12]{2}$$
We thus have
$$\Delta\nu = \nu_0\left(\sqrt[12]{2} - 1\right) \approx 0.0595 \cdot \nu_0 = 0.0595 \cdot 440\,\mathrm{Hz} \approx 26\,\mathrm{Hz}$$
The corresponding time is
$$\Delta t \gtrsim \frac{1}{\Delta\nu} = \frac{1}{26\,\mathrm{Hz}} \approx 38\,\mathrm{ms}$$

- This result comes from a general relationship between time and frequency in periodic quantities—we have to listen long enough to be able to detect frequencies. This should make sense—if we listen less time than the signal's period, we have no way to determine the frequency. The longer we listen to a periodic signal, the more accurately we can determine the frequency.
A similar uncertainty relation holds between any two quantities related by a Fourier transform—for example position and momentum in quantum mechanics.

- Lesson: in practice, use
$$\Delta\nu\Delta t \gtrsim 1$$
when measuring periodic signals.

2.2 Theory: Dependent Measurements

- Consider a physical quantity x which we wish to measure.

  Suppose one lab measures $(\overline{z}_a, \sigma_a^2)$ (average, variance), and a second lab measures $(\overline{z}_b, \sigma_b^2)$, where $\overline{z}_\alpha$ and σ_α^2 denote the average and variance of each set of measurements.

- Throughout this course will assume measurents $\overline{z}_\alpha$ of a quantity x are normally distributed according to

$$\overline{z}_a \sim \mathcal{N}(x, \sigma_a^2)$$

  Some alternate notations for the mean value x are

$$\mathrm{E}\left[\overline{z}_a\right] = \langle \overline{z}_a \rangle = x$$

  The variance σ_a^2 is defined as

$$\mathrm{Var}\left[\overline{z}_a\right] \equiv \left\langle (\overline{z}_a - \langle \overline{z}_a \rangle)^2 \right\rangle = \left\langle (\overline{z}_a - x)^2 \right\rangle = \sigma_a^2.$$

- We can also write measurements in the form

$$\overline{z}_a = x + r_a$$

  where r is random variable representing noise. The noise r is also normally distributed as

$$r_a \sim \mathcal{N}(0, \sigma_a^2),$$

  and the random noise's mean and variance are

$$\langle r_a \rangle = 0 \qquad \text{and} \qquad \mathrm{Var}[r_a] = \left\langle (r_a - \langle r_a \rangle)^2 \right\rangle = \left\langle r_a^2 \right\rangle = \sigma_a^2.$$

Dependence

- We write measurements in noise notation

$$\overline{z}_a = x + r_a \qquad \text{and} \qquad \overline{z}_b = x + r_b.$$

  Two measurements $\overline{z}_a$ and $\overline{z}_b$ are dependent if their noise can be written as

$$r_b = \alpha r_a + w,$$

  where the term αr_a encodes the dependence between $\overline{z}_a$ and $\overline{z}_b$, while w is an independent term, which may be zero, and is Gauss distributed as $w \sim \mathcal{N}(0, \sigma_w^2)$. In the context of dependent noise, we are interested more in the αr_1 term. Because of the independence (of r_a and w?) we have

$$\langle r_a w \rangle = 0$$

  This is the general definition of independence—two noises are independent if the expected value of their product is zero. So r_a and r_b are independent if

$$\langle r_a r_b \rangle = 0.$$

- We return to $r_b = \alpha r_a + w$, and assume r_a and r_b are dependent.

  Using the above definition of dependence we have

  $$\sigma_b^2 = \langle r_b^2 \rangle = \langle (\alpha r_a + w)^2 \rangle = \langle \alpha^2 r_a^2 + w^2 + 2\alpha r_a w \rangle$$

  This reduces (using the expectation value's linearity) to

  $$\sigma_b^2 = \langle r_b^2 \rangle = \alpha^2 \sigma_a^2 + \sigma_w^2 + 2\alpha \langle r_a w \rangle = \alpha^2 \sigma_a^2 + \sigma_w^2,$$

  where we have used $\langle r_1 w \rangle = 0$. The result is

  $$\sigma_b^2 = \alpha^2 \sigma_a^2 + \sigma_w^2$$

  We then divide through by σ_b^2 and get

  $$1 = \left(\alpha \frac{\sigma_a}{\sigma_b}\right)^2 + \left(\frac{\sigma_w}{\sigma_b}\right)^2$$

- Note that both squared terms are positive and their sum is one, so we can bound each inidividual term in the range $(0, 1)$. We thus have

  $$0 \leq \left(\alpha \frac{\sigma_a}{\sigma_b}\right)^2 \equiv \rho_{ab}^2 \leq 1$$

  where we have defined the *correlation coefficient* ρ_{ab}. We then take the square root of both sides of $\rho_{ab}^2 \leq 1$ to get

  $$\rho_{ab} \in [-1, 1].$$

Special Cases

- When $\rho_{ab} = 0$, then r_a and r_b are uncorrelated. This follows mathematically from $\rho_{ab} = 0 \implies \alpha = 0$ (since $\sigma_a, \sigma_b \neq 0$). In this case $r_b = \alpha r_a + w = 0 + w = w$.
- When $\rho_{ab} = \pm 1$ we have

  $$1 = (\rho_{ab})^2 + \left(\frac{\sigma_w}{\sigma_b}\right)^2 = 1 + \left(\frac{\sigma_w}{\sigma_b}\right)^2 \implies \left(\frac{\sigma_w}{\sigma_b}\right)^2 = 0 \implies \sigma_w = 0$$

  The result $\sigma_w = 0$ means r_a and r_b are completey dependent. This relationship is called correlation for $\rho = 1$ and anticorrelation for $\rho = -1$

Covariance

- In our context, the *covariance* of the random variables r_a and r_b is defined as

  $$\sigma_{ab} = \langle r_a r_b \rangle = \langle r_a(\alpha r_a + w) \rangle = \alpha \langle r_a^2 \rangle + \langle r_a w \rangle = \alpha \sigma_a^2 + 0$$

  Note that covariance has the same units as variance σ_α^2, even though it is written without a squared superscript.

- First we note the relationship

  $$\rho_{ab} = \alpha \frac{\sigma_a}{\sigma_b} \implies \alpha \sigma_a = \rho_{ab} \sigma_b$$

  We then see correlation coefficient ρ_{ab} and covariance σ_{ab} are related by

  $$\sigma_{ab} = \alpha \sigma_a^2 = \alpha \sigma_a \cdot \sigma_a = (\rho_{ab} \sigma_b) \cdot \alpha_a$$

  The desired relation is thus

  $$\sigma_{ab} = \rho_{ab} \sigma_a \sigma_b.$$

2.3 Optimal Combination of Dependent Measurements

- Goal: combine the two estimates of the quantity x to get a "sharper" estimate of x with a lesser variance. We'll call this $(\widehat{x}, \sigma^2)$.

  Definition: Optimal combination means combination such that the result has the lowest possible variance.

- Begin with $\overline{z}_a = x + r_a$ and $\overline{z}_b = x + r_b$, where x is the true value and z are measurements and r is noize. Our goal is to combine these optimally into the quantity $\widehat{x}$ which we'll write

  $$\widehat{x} = x + \widehat{r}.$$

  Note that

  $$\langle \widehat{x} \rangle = x \qquad \text{and} \qquad \left\langle (\widehat{x} - x)^2 \right\rangle = \left\langle \widehat{r}^2 \right\rangle = \widehat{\sigma}^2$$

  Our goal is to find $\widehat{\sigma}^2$. We begin with the known $\overline{z}_a$ and $\overline{z}_b$ and σ_a and σ_b

- Write the optimal combination as a linear combination of the input values of the form, using the ansatz

  $$\widehat{x} = \alpha \overline{z}_a + \beta \overline{z}_b$$

  We require a *linear* combination so that the combined result $\widehat{x}$ will also be normally distributed, and so that $\langle \widehat{x} \rangle = x$.

  We then subsitute in the expressions for z to get

  $$\widehat{x} = \alpha(x + r_a) + \beta(x + r_b) = x(\alpha + \beta) + \alpha r_a + \beta r_b$$

  We then write $\widehat{x} = x + \widehat{r}$ to get

  $$x(\alpha + \beta) + \alpha r_a + \beta r_b = x + \widehat{r}$$

  We then combine the random variables on each side of the equation. The x terms are constant and the r terms are variable.

- We then require that $\alpha + \beta = 1$ and thus $\alpha = 1 - \beta$. This requirement comes from the constant x terms being linearly independent of the r terms so the x must be equal to $\alpha + \beta = 1$.

  Important!! the x terms and the random noise terms $r_{a,b}$ are independent.

- Using $\alpha + \beta = 1$ we get

  $$\widehat{x} = \alpha \overline{z}_a + \beta \overline{z}_b = (1 - \beta)\overline{z}_a + \beta \overline{z}_b = \overline{z}_a + \beta(\overline{z}_b - \overline{z}_a)$$

  The end result is

  $$\widehat{x} = \overline{z}_a + \beta(\overline{z}_b - \overline{z}_a).$$

  Our goal is to find the particular b_{opt} such that $\widehat{\sigma}^2$ is minimized.

- To find the optimal b_{opt}, first combine $\widehat{x} = \alpha \overline{z}_a + \beta \overline{z}_b$ with $\widehat{x} = x + \widehat{r}$ to get

  $$\widehat{r} = \alpha r_a + \beta r_b = (1 - \beta) r_a + \beta r_b$$

  Then by definition

  $$\widehat{\sigma}^2 = \left\langle (\widehat{r})^2 \right\rangle = \left\langle [(1 - \beta) r_a + \beta r_b]^2 \right\rangle = (1 - \beta)^2 \sigma_a^2 + \beta^2 \sigma_b^2 + 2\beta(1 - \beta) \langle r_a r_b \rangle$$

  Substituting in $\sigma_{ab} = \langle r_a r_b \rangle$, the optimal variance $\widehat{\sigma}^2$ then reads

  $$\widehat{\sigma}^2 = (1 - \beta)^2 \sigma_a^2 + \beta^2 \sigma_b^2 + 2\beta(1 - \beta)\sigma_{ab}.$$

- We find the optimal coefficient β with basic calculus—we search for the β satisfying

$$\frac{\partial \widehat{\sigma}^2}{\partial \beta} = 0$$

The derivative is

$$\frac{\partial \widehat{\sigma}^2}{\partial \beta} = -2(1-\beta)\sigma_a^2 + 2\beta\sigma_b^2 + 2(1-2\beta)\sigma_{ab} \equiv 0$$

We cancel like terms and rearrange to get

$$\beta(\sigma_a^2 + \sigma_b^2 - 2\sigma_{ab}) = \sigma_a^2 - \sigma_{ab}$$

The optimal value of the coefficient β is thus

$$\beta_{\text{opt}} = \frac{\sigma_a^2 - \sigma_{ab}}{\sigma_a^2 + \sigma_b^2 - 2\sigma_{ab}}.$$

- So β_{opt} substituted into $\widehat{x}$ to get the desired optimal combination $\widehat{x}$. This reads

$$\widehat{x} = \overline{z}_a + \frac{\sigma_a^2 - \sigma_{ab}}{\sigma_a^2 + \sigma_b^2 - 2\sigma_{ab}} \cdot (\overline{z}_b - \overline{z}_a)$$

Some vocabulary $\overline{z}_b - \overline{z}_a$ is called the innovation. The term $\frac{\sigma_a^2 - \sigma_{ab}}{\sigma_a^2 + \sigma_b^2 - 2\sigma_{ab}}$ is called amplification.

Variance of Optimal Estimate

- First an auxiliary calculation

$$1 - \beta_{\text{opt}} = 1 - \frac{\sigma_a^2 - \sigma_{ab}}{\sigma_a^2 + \sigma_b^2 - 2\sigma_{ab}} = \frac{\sigma_b^2 - \sigma_{ab}}{\sigma_a^2 + \sigma_b^2 - 2\sigma_{ab}} \equiv \frac{\sigma_b^2 - \sigma_{ab}}{I}$$

For clarity, in one place

$$\beta_{\text{opt}} = \frac{\sigma_a^2 - \alpha_{ab}}{I} \qquad 1 - \beta_{\text{opt}} = \frac{\sigma_b^2 - \alpha_{ab}}{I}, \quad \text{where } \; I \equiv \sigma_a^2 + \sigma_b^2 - 2\sigma_{ab}$$

- We now find variance of $\widehat{x}$. This is

$$\widehat{\sigma}^2 = (1 - \beta_{\text{opt}})^2\sigma_a^2 + \beta_{\text{opt}}^2\sigma_b^2 + 2\beta_{\text{opt}}(1 - \beta_{\text{opt}})\sigma_{ab}$$

Substitute auxiliary calculations and get

$$\widehat{\sigma}^2 = \left(\frac{\sigma_b^2 - \alpha_{ab}}{I}\right)^2 \sigma_a^2 + \left(\frac{\sigma_a^2 - \alpha_{ab}}{I}\right)^2 \sigma_b^2 + \frac{1}{I^2} \cdot 2\left(\sigma_a^2 - \sigma_{ab}\right)\left(\sigma_b^2 - \sigma_{ab}\right)\sigma_{ab}$$

Factor out the I^2 and get

$$\begin{aligned}\widehat{\sigma}^2 = \frac{1}{I^2}\Big\{&(\sigma_a^2 - \sigma_{ab})\big[(\sigma_a^2 - \sigma_{ab})\sigma_b^2 + (\sigma_b^2 - \sigma_{ab})\sigma_{ab}\big] \\ &+ (\sigma_b^2 - \sigma_{ab})\big[(\sigma_b^2 - \sigma_{ab})\sigma_a^2 + (\sigma_a^2 - \sigma_{ab})\sigma_{ab}\big]\Big\}\end{aligned}$$

After multiplying out the terms in the square brackets and simplifying, this becomes

$$\begin{aligned}\widehat{\sigma}^2 &= \frac{1}{I^2}\Big[(\sigma_a^2-\sigma_{ab})(\sigma_a^2\sigma_b^2-\sigma_{ab}^2)+(\sigma_b^2-\sigma_{ab})(\sigma_a^2\sigma_b^2-\sigma_{ab}^2)\Big]\\ &= \frac{1}{I^2}(\sigma_a^2\sigma_b^2-\sigma_{ab}^2)\big[(\sigma_a^2-\sigma_{ab})+(\sigma_a^2-\sigma_{ab})\big]\\ &= \frac{1}{I^2}(\sigma_a^2\sigma_b^2-\sigma_{ab}^2)(\sigma_a^2+\sigma_b^2-2\sigma_{ab})\end{aligned}$$

Recognizing $(\sigma_a^2+\sigma_b^2-2\sigma_{ab})=I$, cancelling, and using $\sigma_{ab}=\rho_{ab}\sigma_a\sigma_b$ we then get

$$\begin{aligned}\widehat{\sigma}^2 &= \frac{1}{I}\left(\sigma_a^2\sigma_b^2-\sigma_{ab}^2\right)=\frac{1}{I}\left(\sigma_a^2\sigma_b^2-\rho_{ab}^2\sigma_a^2\sigma_b^2\right)=\frac{\sigma_a^2\sigma_b^2}{I}(1-\rho_{ab}^2)\\ &= (1-\rho_{ab}^2)\cdot\frac{\sigma_a^2\sigma_b^2}{\sigma_a^2+\sigma_b^2-2\sigma_{ab}}\end{aligned}$$

Typical notation is

$$\widehat{\sigma}^2=(1-\rho_{ab}^2)\left(\frac{1}{\sigma_a^2}+\frac{1}{\sigma_b^2}-\frac{2\rho_{ab}}{\sigma_a\sigma_b}\right)^{-1}. \tag{2.1}$$

Just to review what this is: it is the variance of $\widehat{x}$, which is the optimally-combined estimate for x, derived from the two *dependent* measurements $\overline{z}_a$ and $\overline{z}_b$, each with variance σ_a and σ_b and covariance

$$\langle(\overline{z}_a-x)(\overline{z}_b-x)\rangle=\sigma_{ab}\neq 0$$

3 Third Exercise Set

3.1 Combining Two Quantities with Equal Variance

- Suppose we have two estimates of a quantity x with known and equal dispersions $\sigma_a = \sigma_b = \sigma$ and an arbitrary correlation coefficient ρ_{ab}.
- We begin with Equation 2.1 for the variance of the optimal combination, which reads

$$\widehat{\sigma}^2_{\text{opt}} = (1-\rho^2_{ab})\left(\frac{1}{\sigma^2_a} + \frac{1}{\sigma^2_b} - \frac{2\rho_{ab}}{\sigma_a\sigma_b}\right)^{-1} \tag{3.1}$$

  and

$$\widehat{x}_{\text{opt}} = \overline{z}_a + \frac{\sigma^2_a - \sigma_{ab}}{\sigma^2_a + \sigma^2_b - 2\sigma_{ab}} \cdot (\overline{z}_b - \overline{z}_a) \tag{3.2}$$

- First we define the *amplification factor*

$$K \equiv \frac{\sigma^2_a - \sigma_{ab}}{\sigma^2_a + \sigma^2_b - 2\sigma_{ab}} = \frac{\sigma^2_a - \sigma_{ab}}{\sigma^2_a + \sigma^2_b - 2\rho_{ab}\sigma_a\sigma_b}.$$

  In our case of equal variances, with $\sigma_a = \sigma_b$, the amplification factor reduces to

$$K = \frac{\sigma^2 - \rho_{ab}\sigma^2}{2(\sigma^2 - \rho_{ab}\sigma^2)} = \frac{1-\rho_{ab}}{2(1-\rho_{ab})} = \frac{1}{2}.$$

  Lesson: the amplification factor for two estimates with equal variances is $K = 1/2$.

- Using $K = 2$ in Equation 3.2 reads, the optimal combination x_{opt} of two equal-variance estimates is

$$\widehat{x}_{\text{opt}} = \overline{z}_a + K \cdot \left(\overline{z}^2 - \overline{z}_a\right) = \overline{z}_a + \frac{1}{2}(\overline{z}_b - \overline{z}_a) = \frac{\overline{z}_a - \overline{z}_b}{2}$$

  Lesson: the optimal combination of the two equal-variance esimates is simply the average of the measurements. In other words, averaging is the optimal form of combination for measurements with *equal variances*.

3.2 Variance of the Mean

- Next, we consider an arbitrary number N of esimates, like the above $\overline{z}_a$. The above formula is defined only for two estimates; in practice we have many estimates. What to do then?
- Next, we consider an arbitrary number N of (in general) dependent measurements $\{z_i\}_{i=1}^N$, distributed normally as

$$z_i \sim \mathcal{N}(x, \sigma^2_i),$$

  where $\sigma_{ij} \neq 0$.

- We define the *mean* of the N measurements with the familiar formula

$$\overline{z} = \frac{1}{N}\sum_i z_i.$$

Using $\langle z_i \rangle = x$, the average $\overline{z}$'s expectation value is

$$\mathrm{E}\big[\overline{z}\big] = \langle \overline{z} \rangle = \left\langle \frac{1}{N} \sum_i z_i \right\rangle = \frac{1}{N} \sum_{i=1}^{N} \langle z_i \rangle = \frac{1}{N} \sum_{i=1}^{N} x = x$$

Lesson: the expected value of the average of N measurements is still the quanity x being measured.

- We now aim to find the variance of the mean, which we write as

$$\sigma_{\mathrm{m}}^2 = \Big\langle (\overline{z} - \langle \overline{z} \rangle)^2 \Big\rangle = \Big\langle (\overline{z} - x)^2 \Big\rangle = \left\langle \left(\frac{1}{N} \sum_{i=1}^{N} z_i - x \right)^2 \right\rangle$$

where we have used the just-derived identity $\mathrm{E}\big[\overline{z}\big] = \langle \overline{z} \rangle = x$. The coefficient $1/N$ is constant; we factor it out and get

$$\sigma_{\mathrm{m}}^2 = \frac{1}{N^2} \left\langle \left(\sum_{i=1}^{N} z_i - Nx \right)^2 \right\rangle = \frac{1}{N^2} \left\langle \left(\sum_{i=1}^{N} (z_i - x) \right)^2 \right\rangle$$

- We then expand the squared term into diagonal and mixed terms. In the mixed terms, we don't sum over $i = j$ because that's accounted for in the diagonal term:

$$\sigma_{\mathrm{m}}^2 = \frac{1}{N^2} \left\langle \left[\sum_{i=1}^{N} (z_i - x)^2 + \sum \sum_{i \neq j} (z_i - x)(z_j - x) \right] \right\rangle .$$

Simplification: in the mixed sum we have double terms: e.g. $i = 1, j = 2$ and $j = 1, i = 2$. Notating that (i, j) terms equal (j, i) terms, we can rewrite the sum as

$$\sigma_{\mathrm{m}}^2 = \frac{1}{N^2} \left\langle \left[\sum_{i} (z_i - x)^2 + 2 \sum_{i<j} (z_i - x)(z_j - x) \right] \right\rangle .$$

Then substitute in known variances of measurements and covariances of different measurements

$$\sigma_{\mathrm{m}}^2 = \frac{1}{N^2} \left(\sum_i \sigma_i^2 + 2 \sum_{i<j}^{N} \sigma_{ij} \right) .$$

Lesson: This is how to find the variance of a mean $\overline{z}$ of N dependent measurements $\{z_i\}$; the mean itself being given by

$$\overline{z} = \frac{1}{N} \sum_{i=1}^{N} z_i .$$

Some Special Cases

- If the measurements $\{z_i\}$ are independent, i.e. if $\sigma_{ij} = 0$ for $i \neq j$, the variance of the mean $\overline{z}$ simplifies to

$$\sigma_{\mathrm{m}}^2 = \frac{1}{N^2} \sum_{i=1}^{N} \sigma_i^2 \tag{3.3}$$

- If the measurements $\{z_i\}$ are all independent with equal variance $\sigma_i \equiv \sigma$, the variance of the mean simplifies further to

$$\sigma_{\mathrm{m}}^2 = \frac{1}{N^2}(N\sigma^2) = \frac{\sigma^2}{N} \implies \sigma_{\mathrm{m}} = \frac{\sigma}{\sqrt{N}}.$$

Lesson: the dispersion of the mean of N independent equal-variance measurements $\{z_i\}$ falls with the square root of the number of measurements.

Generally, we can assume variances of measurements are equal when we measure the quantity x with the same instrument or sensor.

3.3 Example

A GPS receiver at the top of a mountain twice measures the height about sea level and produces two measurements $h_1 = (2139 \pm 12)\,\mathrm{m}$ *and* $h_2 = (2130 \pm 6)\,\mathrm{m}$. *The measurements are uncorrellated.*

1. *Find the average of the measurements and the associated error.*
2. *Find the optimal combination of the two measurements the associated uncertainty. How does the average compare to the optimcal combination?*

- The average of the two measurements is

$$\overline{h} = \frac{h_1 + h_2}{2} = 2134.5\,\mathrm{m}.$$

The average of the mean, using Equation 3.3, is

$$\sigma_{\mathrm{m}}^2 = \frac{1}{N^2}\sum_{i=1}^{N}\sigma_i^2 = \frac{1}{4}\left(12^2 + 6^2\right)\mathrm{m}^2 = 45\,\mathrm{m}^2.$$

The estimate for the average's error is thus

$$\sigma_{\mathrm{m}} \approx 6.7\,\mathrm{m}$$

Note that the error of the average is larger than the error $\sigma_1 = 6\,\mathrm{m}$ of the second measurement. In other words, the process of raw averaging actually *worsened* our estimate (increased the uncertainty) of height.

- We now optimally combine the two measurements. Using Equation 3.2 with $\sigma_{12} = 0$ for independent measurements, the optimal combination $\widehat{h}$ is

$$\widehat{h} = h_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}(h_2 - h_1)$$

The optimal estimate's variance, using Equation 3.1, with $\sigma_{12} = 0$ and $\rho_{12} = 0$, is

$$\widehat{\sigma}^2 = 1 \cdot \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)^{-1} = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2}.$$

Note that in terms of the optimal variance $\widehat{\sigma^2}$, the optimal estimate $\widehat{h}$ may be written

$$\widehat{h} = h_1 + \frac{\widehat{\sigma}^2}{\sigma^2}(h_2 - h_1).$$

- We now compute numerical the results. The optimal estimate is is

$$\widehat{h} = 2139\,\mathrm{m} + \frac{28.8\,\mathrm{m}^2}{(6\,\mathrm{m})^2}(2130 - 2139)\mathrm{m} \approx 2132\,\mathrm{m},$$

while the optimal variance is

$$\widehat{\sigma}^2 = \frac{12^2 \cdot 6^2}{12^2 + 6^2}\mathrm{m}^2 = 28.8\,\mathrm{m}^2 \qquad \text{and} \qquad \widehat{\sigma} \approx 5.4\,\mathrm{m}.$$

The key takeaway here is that $\widehat{\sigma}$ is less than the standard deviation of either of the individual measurements.

3.4 Theory: Kalman Filter for a Constant Quantity

- We measure a constant quantity x with measurements z_i of the form

$$z_i = x + r_i.$$

We call the r_i the measurement noise, and assume this noise uncorrelated, which translates mathematcally to the condition

$$\langle r_i r_j \rangle = \sigma_i^2 \delta_{ij}.$$

- We have a continuous stream of measurements z_i. Our goal is to continuously compute the current optimal estimate of x. We first make some assumptions:
  - After n measurements, we assume we have an optimal estimate $\widehat{x}_n$ and $\widehat{\sigma}_n^2$.
  - At the time $n+1$ we obtain a new measurement z_{n+1} and variance σ_{n+1}^2

Our goal is to optimally combine the previous optimal estimate $(\widehat{x}_n, \widehat{\sigma}_n^2)$ and the new measurement $(z_{n+1}, \sigma_{n+1}^2)$ into an improved optimal estimate $(\widehat{x}_{n+1}, \widehat{\sigma}_{n+1}^2)$. Without proof, this optimal is

$$\widehat{x}_{n+1} = \widehat{x}_n + \frac{\widehat{\sigma}_n^2}{\widehat{\sigma}_n^2 + \sigma_{n+1}^2}(z_{n+1} - \widehat{x}_n) \tag{3.4}$$

The corresponding optimal variance is

$$\widehat{\sigma}_{n+1}^2 = \frac{\widehat{\sigma}_n^2 \sigma^2}{\widehat{\sigma}_n^2 + \sigma^2} \qquad \text{or} \qquad \widehat{\sigma}_{n+1}^{-2} = \widehat{\sigma}_n^{-2} + \sigma_{n+1}^{-2}. \tag{3.5}$$

Equations 3.4 and 3.5 constitute the *Kalman filter* for measuring an unknown, constant, scalar quantity x.

Initializing the Filter

- In the above discussion we started with $\widehat{x}_n$. But how do we start the filter algorithm?

- We begin with
$$\widehat{x}_1 = \widehat{x}_0 + \frac{\widehat{\sigma}_0^2}{\widehat{\sigma}_0^2 + \sigma_1^2}(z_1 - \widehat{x}_0)$$
But we do not know $\widehat{x}_0$. In fact, for all we know its value could be arbitrary!

  Next, since $\widehat{x}_0$ could be anything, we assume corresponding variance is $\widehat{\sigma}_0^2 \to \infty$. An infinite variance is necessary if $\widehat{x}_0$ could assume any value. We then have, applying the limit $\widehat{\sigma}_0^2 \to \infty$, the initial value
$$\widehat{x}_1 = \widehat{x}_0 + 1 \cdot (z_1 - \widehat{x}_0) = z_1 \implies \widehat{x}_1 = z_1.$$
  The corresponding initial variance, from Equation 3.5, is
$$\widehat{\sigma}_1^2 = \lim_{\widehat{\sigma}_0^2 \to \infty} \left(\widehat{\sigma}_0^{-2} + \sigma_1^{-2}\right)^{-1} = \sigma_1^2$$

- The limit approach with $\widehat{\sigma}_0^2 \to \infty$ is evidently equivalent to initializing the filter as $(\widehat{x}_1, \widehat{\sigma}_1^2) = (z_1, \sigma_1^2)$. The lesson is:

  > We initialize the Kalman filter for measuing a constant quantity x using the initial values $(\widehat{x}_1, \widehat{\sigma}_1^2) = (z_1, \sigma_1^2)$, where z_1 is the initial measurement and σ_1^2 is the intial measurement's variance.

3.5 Variance-Weighted Combination

- The Kalman filter combines measurements serially, and is useful in situations in which a continuous stream of measurements arives sequentially one after the other.
- Suppose instead that we are given N measurements all at once, in one bunch. In this case, rather than proceeding one by one as in the Kalman filter, we may optimally combine measurements with "weighted averaging".
- **TODO:** derivation of the below equations is online.
- The optimal combination $\widehat{h}$ of N measurements $\{h_i\}_{i=1}^N$ with variances $\{\sigma_i^2\}_{i=1}^N$ is
$$\widehat{h} = \frac{\sum_{i=1}^N \frac{h_i}{\sigma_i^2}}{\sum_{i=1}^N \sigma_i^{-2}} \equiv \frac{\sum_i w_i h_i}{\sum_i w_i}, \text{ where } w_i \equiv \sigma_i^{-2}.$$
  In the last equality, for shorthand, we have defined the *variance weights* w_i.
- The variance $\widehat{\sigma}^2$ of the above optimal combination $\widehat{h}$ is
$$\widehat{\sigma}^{-2} = \sum_{i=1}^N \sigma_i^{-2} = \sum_{i=1}^N w_i.$$
  See https://en.wikipedia.org/wiki/Weighted_arithmetic_mean#Variance_weights for more on this combination technique.

Homework! Take the equations for $\widehat{h}$ and $\widehat{\sigma}^2$, and use them to derive the Kalman filter equations. In other words, show the above equations for $\widehat{h}$ and $\widehat{\sigma}^2$ give an equivalent result to Kalman filter.

3.6 Theory: Error Propagation with Two Dependent Variables

- Suppose we have the relation $u = f(x, y)$ and we have multiple measurements of the variables: $(\overline{x}, \sigma_{\overline{x}}^2)$ and $(\overline{y}, \sigma_{\overline{y}}^2)$. Suppose the above averages and variances are known.

- Our goal is to find $\sigma_{\overline{u}}^2$. We will do this with the definition of variance:

$$\sigma_{\overline{u}}^2 = \left\langle (\overline{u} - u)^2 \right\rangle$$

  where we assume $\langle \overline{u} \rangle = u$ I think?.

  Using $u = f(x, y)$, this variance reads

$$\sigma_{\overline{u}}^2 = \left\langle \left(\overline{f(x,y)} - f(x,y) \right)^2 \right\rangle$$

  We will then make the Taylor expansion of $f(x, y)$ about the measured values (assume small errors)

$$f(x,y) = f(\overline{x}, \overline{y}) + \left(\frac{\partial f}{\partial x}\right)_{\overline{x},\overline{y}} (x - \overline{x}) + \left(\frac{\partial f}{\partial y}\right)_{\overline{x},\overline{y}} (y - \overline{y}) + \mathcal{O}(2)$$

- Now we find the average value of $f(x, y)$, i.e. $\overline{f(x,y)}$. This is just the equivalent of averaging the expression of $f(x, y)$. This is

$$\overline{f(x,y)} = f(\overline{x}, \overline{y}) + 0 + 0$$

  The first value $f(\overline{x}, \overline{y})$ and the partial derivatives are scalars, while the averages of the differences are zero

- Subsitute $\overline{f(x,y)}$ and $f(x, y)$ into variance. Dropping the $\mathcal{O}(2)$ we have

$$\begin{aligned}
\sigma_{\overline{u}}^2 &\approx \left\langle \left(f(\overline{x}, \overline{y}) - \left[f(\overline{x}, \overline{y}) + \left(\frac{\partial f}{\partial x}\right)_{\overline{x},\overline{y}} (x - \overline{x}) + \left(\frac{\partial f}{\partial y}\right)_{\overline{x},\overline{y}} (y - \overline{y}) \right] \right)^2 \right\rangle \\
&= \left\langle \left(\left(\frac{\partial f}{\partial x}\right)_{\overline{x},\overline{y}} (x - \overline{x}) + \left(\frac{\partial f}{\partial y}\right)_{\overline{x},\overline{y}} (y - \overline{y}) \right)^2 \right\rangle
\end{aligned}$$

  After multiplying out, this reads

$$\begin{aligned}
\sigma_{\overline{u}}^2 \approx &\left(\frac{\partial f}{\partial x}\right)_{\overline{x},\overline{y}}^2 \left\langle (x - \overline{x})^2 \right\rangle + \left(\frac{\partial f}{\partial y}\right)_{\overline{x},\overline{y}}^2 \left\langle (x - \overline{x})^2 \right\rangle \\
&+ 2 \left(\frac{\partial f}{\partial x}\right)_{\overline{x},\overline{y}} \left(\frac{\partial f}{\partial x}\right)_{\overline{x},\overline{y}} \langle (x - \overline{x})(y - \overline{y}) \rangle
\end{aligned}$$

  After evaluating the expectation values, the result is

$$\sigma_{\overline{u}}^2 = \left(\frac{\partial f}{\partial x}\right)_{\overline{x},\overline{y}}^2 \sigma_{\overline{x}}^2 + \left(\frac{\partial f}{\partial y}\right)_{\overline{x},\overline{y}}^2 \sigma_{\overline{y}}^2 + 2 \left(\frac{\partial f}{\partial x}\right)_{\overline{x},\overline{y}} \left(\frac{\partial f}{\partial x}\right)_{\overline{x},\overline{y}} \sigma_{\overline{xy}}.$$

4 Fourth Exercise Set

4.1 Simple Cases of Error Propagation

4.1.1 Linear Sum of Two Variables

- We first consider the linear function

$$u = f(x, y) = ax + by,$$

and assume x and y are independent, so that $\rho_{\overline{xy}} = 0$.

Assume we are given $(\overline{x}, \sigma_{\overline{x}}^2)$ and $(\overline{y}, \sigma_{\overline{y}}^2)$ and $\rho_{\overline{xy}} = 0$.

- To first order, the estimate of u is

$$\overline{u} = f(\overline{x}, \overline{y}) = a\overline{x} + b\overline{y}$$

The error is

$$\sigma_{\overline{u}}^2 - \left(\frac{\partial f}{\partial x}\right)^2_{(\overline{x},\overline{y})} \sigma_{\overline{x}}^2 + \left(\frac{\partial f}{\partial y}\right)^2_{(\overline{x},\overline{y})} \sigma_{\overline{y}}^2$$

where derivatives evaluated at $(\overline{x}, \overline{y})$. We evaulaute derivatives and get

$$\sigma_{\overline{u}}^2 = a^2\sigma_{\overline{x}}^2 + b^2\sigma_{\overline{y}}^2.$$

Lesson: Error propagates in squares. In other words, variances add.

4.1.2 Powers

- We now consider the product of two variables of the form

$$u = f(x, y) = Ax^a y^b$$

We again assume x and y are independent, and have known estimated values and variances $(\overline{x}, \sigma_{\overline{x}}^2)$ and $(\overline{y}, \sigma_{\overline{y}}^2)$.

- The estimate for u using the estimates $\overline{x}$ and $\overline{y}$ is

$$\overline{u} = A\overline{x}^a \cdot \overline{y}^b.$$

To find the corresponding variance, we first make the auxiliar calculations

$$\frac{\partial f}{\partial x} = Aax^{a-1}y^b \implies \left(\frac{\partial f}{\partial x}\right)^2_{(\overline{x},\overline{y})} = A^2a^2\overline{x}^{2(a-1)}\overline{y}^{2b},$$

and

$$\frac{\partial f}{\partial y} = Abx^{\delta}y^{b-1} \implies \left(\frac{\partial f}{\partial y}\right)^2_{(\overline{x},\overline{y})} = A^2b^2\overline{x}^{2a}\overline{y}^{2(b-1)}.$$

Using these derivatives, the variance of the estimate $\overline{u}$ is

$$\sigma_{\overline{u}}^2 = A^2a^2\overline{x}^{2a}\overline{x}^{-2}\overline{y}^{2b}\sigma_{\overline{x}}^2 + A^2b^2\overline{x}^{2a}\overline{y}^{2\phi}\overline{y}^{-2}\sigma_{\overline{y}}^2.$$

- In this case, for products the product of two variables, a useful quantity is relative error. We first find

$$\overline{u}^2 = A^2 \overline{x}^{2a} \overline{y}^{2b}$$

In this case relative error $\sigma_{\overline{u}}/\overline{u}$ (squared) is fairly simple:

$$\left(\frac{\sigma_{\overline{u}}}{\overline{u}}\right)^2 = a^2 \left(\frac{\sigma_{\overline{x}}}{\overline{x}}\right)^2 + b^2 \left(\frac{\sigma_{\overline{y}}}{\overline{y}}\right)^2,$$

where we have recognized the expression for the relative errors of the input quantities x and y.

4.2 Theory: The Normal (Gaussian) Distribution

- Let a random variable z be distributed with mean μ and variance σ^2 as

$$z \sim p(z) = \frac{\mathrm{d}P}{\mathrm{d}z} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-\mu)^2}{2\sigma^2}},$$

which we write in shorthand as

$$z \sim \mathcal{N}(\mu, \sigma^2)$$

- Typically we are interested in the probability P that the variable z is less than some value x. Write this as

$$P(z \leq x) = \int_{-\infty}^{x} \frac{\mathrm{d}P}{\mathrm{d}z}\,\mathrm{d}z = F(x; \mu, \sigma^2),$$

where we have defined the cumulative distribution function (CDF) F. This CDF is not easily calculated in terms of analytic functions. More so, in principle, it would have to calculated anew for each pair (μ, σ^2).

- The solution is the transformation to a standardized Gaussian distribution. We first define the new random variable

$$u \equiv \frac{z-\mu}{\sigma} \implies \mathrm{d}z = \sigma\,\mathrm{d}u.$$

This new variable u is distributed as

$$u \sim p(u) = \frac{\mathrm{d}P}{\mathrm{d}u} = \frac{\mathrm{d}P}{\mathrm{d}z}\frac{\mathrm{d}z}{\mathrm{d}u} = \frac{\mathrm{d}P}{\mathrm{d}z} \cdot \sigma.$$

In this case the distribution is the *standardized normal distribution*

$$u \sim \frac{\mathrm{d}P}{\mathrm{d}u} = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} = \mathcal{N}(0, 1).$$

- Why is this useful? The standardized normal distribution allows us to transform an arbitrary normal distribution into a standard form, which allows us to determine a single CDF we can use in general. The standard Gaussian CDF is

$$\Phi(x) = \int_{-\infty}^{x} \frac{\mathrm{d}P}{\mathrm{d}u}\,\mathrm{d}u.$$

As before, we still must compute $\Phi(x)$ numerically. But we only need to do this for a single CDF. In practice, we simply tabulate values of $\Phi(x)$ for typical values of x.

- Keep in mind the identities, for reference

$$\Phi(-\infty) = 0 \qquad \Phi(0) = \frac{1}{2} \qquad \Phi(\infty) = 1 \qquad \Phi(-x) = 1 - \Phi(x).$$

The last identity is particularly useful—it means we only need to tabulate values of $\Phi(x)$ only for e.g. positive x, and use $\Phi(-x) = 1-\Phi(x)$ to compute corresponding values of $\Phi(-x)$.

- Next, we consider the probability for finding a random variable $z \sim \mathcal{N}(\mu, \sigma^2)$ in the region e.g. (α, β), which we write as $P(\alpha < z < \beta)$. This propability can be written

$$P(\alpha < z < \beta) = P(z < \beta) - P(z < \alpha) = F(\beta) - F(\alpha).$$

- It is more computationally practical to define this probability in terms of the standardized variable u, defined as

$$u = \frac{z - \mu}{\sigma}$$

The values of α and β transform into

$$u_\alpha = \frac{\alpha - \mu}{\sigma} \qquad \text{and} \qquad u_\beta = \frac{\beta - \mu}{\sigma}.$$

The probability then becomes

$$\begin{aligned} P(\alpha < z < \beta) &= P(u_\alpha < u < u_\beta) = P(u < u_\beta) - P(u < u_\alpha) \\ &= \Phi(u_\beta) - \Phi(u_\alpha) \\ &= \Phi\Big(\tfrac{\beta-\mu}{\sigma}\Big) - \Phi\Big(\tfrac{\alpha-\mu}{\sigma}\Big) \end{aligned}$$

- Next, we consider a special case, if $\alpha = \mu - \sigma$ and $\beta = \mu + \sigma$, i.e. if the limits are placed symmetrically one standard deviation away from the mean value μ. In this case we have

$$P(\mu - \sigma < z < \mu + \sigma) = \Phi(1) - \Phi(-1) = f(1) - \big[1 - \Phi(1)\big] = 2\Phi(1) - 1.$$

The value of $\Phi(1)$ may be found from a standard table; it is $\Phi(1) \approx 0.8413$. The above probability is then

$$P(\mu - \sigma < z < \mu + \sigma) = 2 \cdot 0.8413 - 1 \approx 0.68 \approx \frac{2}{3}.$$

Finally we note the identity

$$P(\mu - n\sigma < z < \mu + n\sigma) = 2\Phi(n) - 1.$$

x	0	1	2	3	4	5	6	7	8	9
0.0	0.5000	5040	5080	5120	5160	5199	5239	5279	5319	5359
0.1	5398	5438	5478	5517	5557	5596	5636	5675	5714	5753
0.2	5793	5832	5871	5910	5948	5987	6026	6064	6103	6141
0.3	6179	6217	6255	6293	6331	6368	6406	6443	6480	6517
0.4	6554	6591	6628	6664	6700	6736	6772	6808	6844	6879
0.5	6915	6950	6985	7019	7054	7088	7123	7157	7190	7224
0.6	7257	7291	7324	7357	7389	7422	7454	7486	7517	7549
0.7	7580	7611	7642	7673	7704	7734	7764	7794	7823	7852
0.8	7881	7910	7939	7967	7995	8023	8051	8078	8106	8133
0.9	8159	8186	8212	8238	8264	8289	8315	8340	8365	8389
1.0	8413	8438	8461	8485	8508	8531	8554	8577	8599	8621
1.1	8643	8665	8686	8708	8729	8749	8770	8790	8810	8830
1.2	8849	8869	8888	8907	8925	8944	8962	8980	8997	9015
1.3	9032	9049	9066	9082	9099	9115	9131	9147	9162	9177
1.4	9192	9207	9222	9236	9251	9265	9279	9292	9306	9319
1.5	9332	9345	9357	9370	9382	9394	9406	9418	9429	9441
1.6	9452	9463	9474	9484	9495	9505	9515	9525	9535	9545
1.7	9554	9564	9573	9582	9591	9599	9608	9616	9625	9633
1.8	9641	9649	9656	9664	9671	9678	9686	9693	9699	9706
1.9	9713	9719	9726	9732	9738	9744	9750	9756	9761	9767
2.0	9772	9778	9783	9788	9793	9798	9803	9808	9812	9817
2.1	9821	9826	9830	9834	9838	9842	9846	9850	9854	9857
2.2	9861	9864	9868	9871	9875	9878	9881	9884	9887	9890
2.3	9893	9896	9898	9901	9904	9906	9909	9911	9913	9916
2.4	9918	9920	9922	9925	9927	9929	9931	9932	9934	9936
2.5	9938	9940	9941	9943	9945	9946	9948	9949	9951	9952
2.6	9953	9955	9956	9957	9959	9960	9961	9962	9963	9964
2.7	9965	9966	9967	9968	9969	9970	9971	9972	9973	9974
2.8	9974	9975	9976	9977	9977	9978	9979	9979	9980	9981
2.9	9981	9982	9982	9983	9984	9984	9985	9985	9986	9986
3.0	9987	9987	9987	9988	9988	9989	9989	9989	9990	9990
3.1	9990	9991	9991	9991	9992	9992	9992	9992	9993	9993
3.2	9993	9993	9994	9994	9994	9994	9994	9995	9995	9995
3.3	9995	9995	9995	9996	9996	9996	9996	9996	9996	9997
3.4	9997	9997	9997	9997	9997	9997	9997	9997	9997	9998
3.5	9998	9998	9998	9998	9998	9998	9998	9998	9998	9998
3.6	9998	9998	9999	9999	9999	9999	9999	9999	9999	9999
3.7	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999
3.8	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999

Table 1: Values of the standard Gaussian cumulative distribution function for values of x in the range $x \in [0, 3.89]$.

4.3 Estimating Reflection Probability with a Normal Distribution

We know the kinetic energy of a particle to an accuracy of four percent. The particle is incident on a potential barrier, which is one percent lower than the particle's kinetic energy. Determine the probability for reflection.

- Note we are given basically expected values.

  Potential barrier height is given by $V_0 = 0.99\overline{T}$.

- We assume the kinetic energy is distributed normally with mean $\overline{T}$ and standard deviation $\sigma_T = 0.04\overline{T}$.

- Reflection occurs if the particle's kinetic energy is less than the height of the potential barrier. In symbols, the reflection probability is written

  $$P_{\mathrm{r}} = P_T(T < V_0) = F(V_0),$$

  where the quantity $P_T(T < V_0)$ is the probability that the particle's kinetic energy T is less than the height of the potential barrier.

- The standardized variable corresponding to T, e.g. $\mathcal{T}$, is

  $$\mathcal{T} = \frac{T - \mu_T}{\sigma_T} = \frac{T - \overline{T}}{0.04 \cdot \overline{T}} \implies \mathcal{T}(V_0) = \frac{V_0 - \overline{T}}{0.04 \cdot \overline{T}} = \frac{0.99 \cdot \overline{T} - \overline{T}}{0.04 \cdot \overline{T}} = -\frac{1}{4}.$$

  Okay then

  $$F(V_0) = \Phi\left(\frac{V_0 - \overline{T}}{0.04 \cdot \overline{T}}\right) = \Phi(\mathcal{T}(V_0)) = \Phi\left(-\frac{1}{4}\right) = 1 - \Phi\left(\frac{1}{4}\right) \approx 1 - 0.6 \approx 0.4.$$

  We expect the probability for reflection is less than 0.5, since the particle's kinetic energy is lower than the height of the barrier.

4.4 Theory: Kalman Filter for a Scalar Variable

- We begin with a review of the Kalman filter for measuring a scalar variable. Assume we are interested in measuring a physical quantity $x = x(t)$, which we assume obeys the first order linear differential equation

  $$\frac{\mathrm{d}x}{\mathrm{d}t} = A(t)x(t) + b(t) \tag{4.1}$$

  This model of x's dynamics is, evidently, continuous. We then transform the continuous model into a discrete model with the difference equation

  $$\frac{x_{n+1} - x_n}{T} = A(nT) \cdot x(nT) + b(nT) \equiv A_n x_n + b_n,$$

  where T is the time between subsequent measurements of the quantity x.

- We then multiply through by T and rearrange to get

  $$x_{n+1} = x_n\big[1 + A(nT) \cdot T\big] + c_n,$$

  where we have defined

  $$\Phi_n \equiv 1 + A_{nT} \cdot T \qquad \text{and} \qquad c_n \equiv b_n \cdot T.$$

  In this notation, the continuous dynamics equation (Eq. 4.1) reads

  $$x_{n+1} = \Phi_n x_n + c_n.$$

- Finally, allowing for dynamic noise, we write the equation as

$$x_{n+1} = \Phi_n x_n + c_n + \Gamma_n w_n,$$

where w_n is dynamic noise with variance $\langle w_n^2 \rangle \equiv Q$, while the Γ_n term is included only for similarity to the dynamics of a vector variable, which we will see a few sections later.

- Next: *assume* that at time $t = Tn$ we have an optimal estimate $\widehat{x}_n$ of the quantity being measured, together with the corresponding variance $\widehat{\sigma}_n^2$. We then read off a new measurement $(z_{n+1}, \sigma_{n+1}^2)$.

- We proceed by mapping the previous optimal estimate $(\widehat{x}_n, \widehat{\sigma}_n^2)$ forward in time to get the extrapolated value $(\overline{x}_{n+1}, \overline{\sigma}_{n+1}^2)$. We make this extrapolation with the assumed-to-be-known dynamics governing the quantity x.

- The next step is to optimize the extrapolated value $(\overline{x}_{n+1}, \overline{\sigma}_{n+1}^2)$ using the measured value $(z_{n+1}, \sigma_{n+1}^2)$ and the known Kalman filter procedure for a scalar constant.

- We first define the variances with the quantities

$$P_n \equiv \widehat{\sigma}_n^2 \qquad M_{n+1} \equiv \overline{\sigma}_{n+1}^2 \qquad R_{n+1} \equiv \sigma_{n+1}^2.$$

The extrapolation step (see lecture notes for a derivation) is given by the equations

$$\begin{aligned} \overline{x}_{n+1} &= \Phi_n \widehat{x}_n + c_n, \\ M_{n+1} &= \Phi_n^2 P_n + \Gamma_n^2 Q_n. \end{aligned}$$

- The second step, i.e. optimizing the extrapolated estimate, reads

$$\begin{aligned} \widehat{x}_{n+1} &= \overline{x}_{n+1} + \frac{M_{n+1}}{M_{n+1} + R_{n+1}}(z_{n+1} - \overline{x}_{n+1}) \equiv \overline{x}_{n+1} + K_{n+1}(z_{n+1} - \overline{x}_{n+1}) \\ P_{n+1} &= \frac{M_{n+1} R_{n+1}}{M_{n+1} + R_{n+1}} = M_{n+1} - \frac{M_{n+1}^2}{M_{n+1} + R_{n+1}}, \end{aligned}$$

where we have define the amplification factor

$$K_{n+1} = \frac{M_{n+1}}{M_{n+1} + R_{n+1}} = \frac{P_{n+1}}{R_{n+1}}.$$

- Finally note that if we are not performing measurements, i.e. if the value of z_{n+1} is arbitrary, then $R_{n+1} \to \infty$ since z_{n+1} could be anything.

The optimal estimate the reduces to $\widehat{x}_{n+1} = \overline{x}_{n+1}$, which makes sense, since we don't have any more information than the extrapolation itself. The corresponding variance of the optimal estimate is then $P_{n+1} = M_{n+1}$.

5 Fifth Exercise Set

5.1 Kalman Filter for a Ball Bouncing Down Stairs

A ball rolls to the top of a staircase of five equally tall stairs with exactly known height $\overline{H} = 30\,\mathrm{cm}$ *and zero uncertainty,* $\sigma_H^2 = 0\,\mathrm{m}$. *The ball rolls down the stairs, bouncing exactly once on each stair. During each bounce, the ball looses half of its kinetic energy associated with vertical motion.*
Determine the maximum height of the final bounce and the associated dispersion. There is no measurement of height during the intermediate bounces. Neglect dynamic noise. The initial height of the ball is zero, but is known only up to the standard deviation σ_0.

- We will solve the problem with the Kalman filter. Our first step is to determine the ball's dynamics, and write the dynamics in the form matching the Kalman filter algorithm, which reads

  $$x_{n+1} = \Phi_n x_n + c_n.$$

  xThe variable of interest is the ball height h; we aim to write the ball's height after the n-th bounce in the form

  $$h_{n+1} = \Phi_n h_n + c_n,$$

  which will allow us to determine Φ_n and c_n.

- Note that we need the Kalman filter only to compute the final uncertainty σ_5. We can find the height with basic mechanics.

- Let $\delta = 1/2$ be the fraction of conserved kinetic energy.

- Let h_0 denote the ball's initial height above the first step. (In our case we have $h_0 = 0\,\mathrm{cm}$, but we will solve in the general case.)

- The ball's potential energy U_1 after the first bounce is

  $$U_1 = \delta \cdot U_0 \implies mgh_1 = \delta \cdot mg(h_0 + H) \implies h_1 = \delta h_0 + \delta H.$$

  Note formally $\overline{H}$ but zero uncertainty so H is cool.

  Continuing the pattern, the potential energy fater the second bounce is

  $$U_2 = \delta \cdot U_1 \implies mgh_2 = \delta \cdot mg(h_1 + H) \implies h_2 = \delta h_1 + \delta H$$

  The height on the RHS is the height of the stair H plus the height h_n above the stair at the top of the n-th bounce.

- Generalizing the pattern, the relationship between heights at successive bounces is

  $$h_{n+1} = \delta \cdot h_n + \delta H.$$

  This is the desired dynamic equation. Compared to the general Kalman filter equation $h_{n+1} = \Phi_n h_n + c_n$, we see $\Phi_n = \delta$ and $c_n = \delta H$.

- Return to dynamic equation

  $$h_{n+1} = \delta h_n + \delta H,$$

  where we have identified $\Phi_n = \delta$ and $c_n = \delta H$.

- Next, turn to computing h_5. We begin by finding a recursive height relation; the goal is to express h_n in terms of h_0. Begin with, for example

$$h_2 = \delta h_1 + \delta\overline{H} = \delta(\delta h_0 + \delta\overline{H}) + \delta\overline{H} = \delta^2 h_0 + \overline{H}(\delta + \delta^2)$$

For h_3, we then have

$$h_3 = \delta h_2 + \delta\overline{H} = \delta\big[\delta^2 h_0 + \overline{H}(\delta + \delta^2)\big] + \delta\overline{H} = \delta^3 h_0 + \overline{H}(\delta + \delta^2 + \delta^3)$$

From here we can identify the general expression for the height h_n above the n-th step. This reads

$$h_n = \delta^n h_0 + \overline{H}\sum_{k=1}^{n}\delta^k.$$

Finally we simplify the sum

$$S_n \equiv \sum_{k=1}^{n}\delta^k = \delta(1 + S_n - \delta^n) \implies S_n = \delta\frac{1-\delta^n}{1-\delta}$$

The maximum height above the n-th step is then

$$h_n = \delta^n h_0 + \overline{H}\cdot\delta\frac{1-\delta^n}{1-\delta}.$$

- For the concrete case in our problem with $n = 5$, $\delta = 1/2$ we have

$$h_5 = \overline{H}\cdot\frac{1}{2}\cdot\frac{1-(1/2)^5}{1-1/2} = \overline{H}\cdot\big(1 - \tfrac{1}{32}\big) = \frac{31}{32}\cdot\overline{H}$$

- Next, we use the Kalman filter to find the variance σ_5^2 of the height h_5.

In our case we only have one step in the Kalman filter, since their are no in-between measurements of height during each bounce. In other words, we can only estimate the height with the ball's known dynamics, but we don't have measurements to improve each estimate.

Writing a hat to match the Kalman filter notation, we have

$$\sigma_5^2 \to \widehat{\sigma}_5^2 = P_5$$

Because we have no in-between measurements, we have $P_{n+1} = M_{n+1}$. This is then

$$P_{n+1} = M_{n+1} = \Phi_n^2 P_n + \underbrace{\Gamma_n^2 Q_n}_{=0} \implies P_{n+1} = \Phi_n^2 P_n.$$

Using the known coefficient $\Phi_n = \delta$, we have

$$P_{n+1} = \delta^2 P_n \qquad \text{and} \qquad P_n = \big(\delta^2\big)^n P_0.$$

- First, we note that P_0 is the initial variance of the ball's height; this is $P_0 = \sigma_0^2$. Our goal is to write P_n in terms of the known P_0. This is simply

$$P_n = \big(\delta^2\big)^n P_0 = \delta^{2n}\sigma_0^2$$

We are interested in dispersion (standard deviation in this case) and not variance. This is

$$\widehat{\sigma}_n = \sqrt{P_n} = \delta^n \cdot \sigma_0.$$

In our case with $\delta = 1/2$ and $n = 5$ we have

$$\widehat{\sigma}_5 = \frac{\sigma_0}{32}.$$

The hat because this expression is our optimal estimate for the fifth height's variance.

Note that we would have an even lower value of $\widehat{\sigma}_5$ if measurements were known during each bounce. Even so, the variance decreases from the initial value because of the problem's dynamics alone.

5.2 Theory: Kalman Theory for Vector Measurements

TODO:

5.3 Kalman Filter for an Elastic Collision

Two bodies with masses m_1 and m_2 and speeds v_1 and v_2 collide elastically. The uncertainties in speed are known and equal to σ_1 and σ_2. Determine the covariance matrix for the vector quantity $\boldsymbol{v}' = \left(v_1', v_2'\right)^\top$ representing the bodies' velocities after the collision. Assume the pre-collision velocities v_1 and v_2 are uncorrelated, and that the bodies' dynamics are exactly known..

- We first define the vector quantities

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{v}' = \begin{pmatrix} v_1' \\ v_2' \end{pmatrix},$$

  which represent the bodies' speeds before and after the collision, respectively. The covariance matrix for the pre-collision velocity vector is

$$\mathbf{P} = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}, \tag{5.1}$$

  where σ_1^2 and σ_2^2 are the variances (given in the instructions) of the pre-collision speeds v_1 and v_2.

  Some notes:

  - For simplicity, we will assume the bodies slide without friction, to avoid the dynamics of rolling.
  - We don't have measurements, so the Kalman filter involves only an extrapolation step.

    Note that hats for optimal estimates are implicit.

- Our goal is to find $\mathbf{P}'$, the velocity covariance matrix after the collision. We will find $\mathbf{P}'$ with the Kalman filter equation

$$\mathbf{P}' = \mathbf{M}' = \boldsymbol{\Phi}\mathbf{P}\boldsymbol{\Phi}^\top + \underbrace{\boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^\top}_{=\mathbf{0}} \implies \mathbf{P}' = \boldsymbol{\Phi}\mathbf{P}\boldsymbol{\Phi}^\top, \tag{5.2}$$

where $\mathbf{M}'$ is the covariance matrix of the extrapolation after the collision, and is equal to $\mathbf{P}'$ because there are no measurements. Since the bodies' dynamics are known exactly, the dynamic noise $\mathbf{Q}$ is zero. Our problem thus reduces to finding the matrix $\mathbf{\Phi}$, which governs the dynamics of the two bodies.

Dynamics

- To find the matrix $\mathbf{\Phi}$, we must write the two-ball system's dynamics in the form

$$\boldsymbol{v}' = \mathbf{\Phi}\boldsymbol{v} + \boldsymbol{c}.$$

  In other words, we aim to write the post-collision velocity $\boldsymbol{v}'$ in terms of the pre-collision velocity $\boldsymbol{v}$; the coefficients of the $\boldsymbol{v}$ terms will then determine $\mathbf{\Phi}$.

- We will solve the collision problem in the center-of-mass (CMS) system, where the dynamics of the collision are simpler, which will make it easier to define $\mathbf{\Phi}$.

  The center of mass speed v^* in the lab frame is given by

$$v^* = \frac{m_1 v_1 + m_2 v_2}{m_1 + m_2}$$

  In the center of mass frame, we will write speeds as u. Before the collision, the CMS speeds of two bodies are

$$u_1 = v_1 - v^* \qquad \text{and} \qquad u_2 = v_2 - v^*.$$

- Assuming an elastic collision, both the system's momentum and kinetic energy are conserved. In the CMS system, the general expressions for conservation of momentum p and kinetic energy T for a collision of n bodies are

$$\sum_i^n p_i = \sum_{i=1}^n m_i u_i = 0 \qquad \text{and} \qquad \sum_{i=1}^n T_i = \sum_{i=1}^n T_i',$$

  where T_i and T_i' are the kinetic energies of the i-th body before and after the collision.

- We begin with conservation of momentum. Before the collision, this reads

$$m_1 u_1 + m_2 u_2 = 0 \implies u_2 = -\frac{m_1}{m_2} u_1,$$

  while after the collision we have

$$m_1 u_1' + m_2 u_2' = 0 \implies u_2' = -\frac{m_1}{m_2} u_1'.$$

- Next, we apply conservation of kinetic energy. After cancelling 1/2 factors, this produces

$$m_1 u_1^2 + m_2 u_2^2 = m_1 (u_1')^2 + m_2 (u_2')^2,$$

  which we combine with the expressions for u_2 and u_2' from momentum conservation to get

$$m_1 u_1^2 + \frac{m_1^2}{m_2} u_1^2 = m_1 (u_1')^2 + \frac{m_1^2}{m_2} (u_1')^2.$$

We then factor and cancel like terms, producing

$$u_1^2\left(m_1+\frac{m_1^2}{m_2}\right)=(u_1')^2\left(m_1+\frac{m_1^2}{m_2}\right)\implies (u_1')^2=u_1^2\implies \boxed{u_1'=\pm u_1}.$$

We choose the solution with a negative sign.[2] The first body's post-collision velocity in the center of mass system is thus $u_1'=-u_1$.

- Deriving the second body's post-collision speed is completely analogous, and we simply quote the result:

$$u_2'=-u_2.$$

Alternatively, we could simply note that all of the above kinematic equations are invariant under change of index. We could then simply change indeces in the above derivation and get $u_2'=-u_2$.

Transformation Back to Lab System

- We now transform the CM speeds back into the lab system using the general relationship $u_i=v_i-v^*$ together with the results $u_i'=-u_i$, and the fact that the center of mass velocity v^* is preserved for an elastic collision.

For example, for the first body we have

$$v_1'-v^*=u_1'=(-u_1)=-(v_1-v^*)\implies v_1'=-v_1+2v^*.$$

Substituting in the center of mass velocity, making a common denominator, and combining like terms produces

$$\begin{aligned}v_1'&=-v_1+2v^*=-v_1+2\frac{m_1v_1+m_2v_2}{m_1+m_2}=\frac{v_1(m_1-m_2)+2m_2v_2}{m_1+m_2}\\&\equiv\frac{v_1(1-\mu)+2\mu v_2}{1+\mu},\end{aligned}$$

where we have defined the constant $\mu\equiv m_2/m_1$.

- An analogous calculation for the second body (or simply applying the index invariance argument discussed above) results in

$$v_2'=\frac{v_2(m_2-m_1)+2m_1v_1}{m_1+m_2}=\frac{v_2(\mu-1)+2v_1}{1+\mu}.$$

Kalman Filter Dynamics

- Recall that we aim to find the matrix $\boldsymbol{\Phi}$ in the Kalman filter equation

$$\boldsymbol{v}'=\boldsymbol{\Phi}\boldsymbol{v}+\boldsymbol{c}.$$

To identify $\boldsymbol{\Phi}$ and $\boldsymbol{c}$, we write the just-derived expressions for v_1' and v_2' in the form

$$\begin{aligned}v_1'&=\frac{1-\mu}{1+\mu}v_1+\frac{2\mu}{1+\mu}v_2\\v_2'&=\frac{2}{1+\mu}v_1+\frac{\mu-1}{1+\mu}v_2.\end{aligned}$$

[2] The solution $u_1'=+u_1$, in which the post-collision speed and direction equal their initial values, would mean no collision at all! The negative sign corresponds to a change of direction after collision.

- From the coefficients of v_1 and v_2, we can then read off the matrix equation

$$\begin{pmatrix} v_1' \\ v_2' \end{pmatrix} = \frac{1}{1+\mu} \begin{pmatrix} 1-\mu & 2\mu \\ 2 & \mu-1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

  Evidently, we have $\boldsymbol{c} = \mathbf{0}$, while the desired dynamics matrix $\boldsymbol{\Phi}$ reads

$$\boldsymbol{\Phi} = \frac{1}{1+\mu} \begin{pmatrix} 1-\mu & 2\mu \\ 2 & \mu-1 \end{pmatrix}.$$

Covariance

- Next, using the just-derived dynamics matrix $\boldsymbol{\Phi}$, we will find the velocity covariance matrix $\mathbf{P}'$ after the collision using Equation 5.2, which for review reads

$$\mathbf{P}' = \boldsymbol{\Phi}\mathbf{P}\boldsymbol{\Phi}^\top.$$

  Using the known velocity covariance matrix $\mathbf{P}$ from Equation 5.1, this comes out to

$$\begin{aligned} \mathbf{P}' = \boldsymbol{\Phi}\mathbf{P}\boldsymbol{\Phi}^\top &= \frac{1}{(1+\mu)^2} \begin{pmatrix} 1-\mu & 2\mu \\ 2 & \mu-1 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \begin{pmatrix} 1-\mu & 2 \\ 2\mu & \mu-1 \end{pmatrix} \\ &= \frac{1}{(1+\mu)^2} \begin{pmatrix} 1-\mu & 2\mu \\ 2 & \mu-1 \end{pmatrix} \begin{pmatrix} \sigma_1^2(1-\mu) & 2\sigma_1^2 \\ 2\mu\sigma_2^2 & \sigma_2^2(\mu-1) \end{pmatrix} \\ &= \frac{1}{(1+\mu)^2} \begin{pmatrix} (1-\mu)^2\sigma_1^2 + 4\mu^2\sigma_2^2 & 2(1-\mu)\sigma_1^2 + 2\mu(\mu-1)\sigma_2^2 \\ 2(1-\mu)\sigma_1^2 + 2\mu(\mu-1)\sigma_2^2 & 4\sigma_1^2 + (\mu-1)^2\sigma_2^2 \end{pmatrix} \end{aligned} \tag{5.3}$$

- *Tip*: when multiplying matrices with complicated matrix elements, write the matrix elements as shorthand constants, and only substitute back at the end. In our case this would read

$$A \equiv \frac{1}{1+\mu} \quad a \equiv 1-\mu \quad b \equiv 2\mu \quad c \equiv 2 \implies \boldsymbol{\Phi} = A \begin{pmatrix} a & b \\ c & -a \end{pmatrix}.$$

  The matrix multiplication the reads

$$\begin{aligned} \mathbf{P}' &= A^2 \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \begin{pmatrix} a & c \\ b & -a \end{pmatrix} \\ &= A^2 \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \begin{pmatrix} a\sigma_1^2 & c\sigma_1^2 \\ b\sigma_2^2 & -a\sigma_2^2 \end{pmatrix} \\ &= A^2 \begin{pmatrix} a^2\sigma_1^2 + b^2\sigma_2^2 & ac\sigma_1^2 - ab\sigma_2^2 \\ ac\sigma_1^2 - ab\sigma_2^2 & c^2\sigma_1^2 + a^2\sigma_2^2 \end{pmatrix} \end{aligned}$$

  After substituting the constants back in, the result is

$$\mathbf{P}' = \frac{1}{(1+\mu)^2} \begin{pmatrix} (1-\mu)^2\sigma_1^2 + 4\mu^2\sigma_2^2 & 2(1-\mu)(\sigma_1^2 - \mu\sigma_2^2) \\ 2(1-\mu)(\sigma_1^2 - \mu\sigma_2^2) & 4\sigma_1^2 + (1-\mu)^2\sigma_2^2 \end{pmatrix}, \tag{5.4}$$

  in agreement with Equation 5.3. Using single-letter constants tends to reduces the risk of arithmetic errors.

  Tip: Covariance matrices are always symmetric. Leverage this symmetry either as a safety check, to ensure you multiplied correctly, or to skip on calculating matrix elements.

- We will now analyze the covariance matrix in Equation 5.4 for a few special cases.

  1. First, we consider the case $\sigma_1 = \sigma_2 \equiv \sigma$, i.e. when the uncertainties in initial speeds are equal. This produces

  $$\mathbf{P}' = \frac{\sigma^2}{(1+\mu)^2}\begin{pmatrix}(1-\mu)^2+4\mu^2 & 2(1-\mu)^2 \\ 2(1-\mu)^2 & 4+(1-\mu)^2\end{pmatrix}.$$

  2. Next, consider the further case $\sigma_1 = \sigma_2$ *and* $m_2 = m_1$, in which case $\mu = 1$. This gives

  $$\mathbf{P}' = \frac{\sigma^2}{4}\begin{pmatrix}4 & 0 \\ 0 & 4\end{pmatrix} = \sigma^2\mathbf{I} = \mathbf{P}.$$

  In other words, variance after the collision equals the variance before the collision.

  3. Finally, we consider the case $\sigma_1 \neq \sigma_2$ and $m_1 = m_2 \implies \mu = 1$. This produces post-collision covariance matrix

  $$\mathbf{P}' = \frac{1}{4}\begin{pmatrix}4\sigma_2^2 & 0 \\ 0 & 4\sigma_1^2\end{pmatrix} = \begin{pmatrix}\sigma_2^2 & 0 \\ 0 & \sigma_1^2.\end{pmatrix}$$

  Recall that the initial, pre-collision covariance matrix reads

  $$\mathbf{P} = \begin{pmatrix}\sigma_1^2 & 0 \\ 0 & \sigma_2^2\end{pmatrix}.$$

  In other words, the role of uncertainties in $\mathbf{P}'$ and $\mathbf{P}$ are switched before and after the collision. Interpretted physically, this occurs because speeds switch in an elastic collision of two equal masses.

Homework

- Consider the same problem, but for a completely inelastic collision. The two bodies collide and stick together; we have $v_1' = v_2'$.
- In this case it works to solve the problem in the lab system.
- Momentum is conserved, while kinetic energy is not conserved.
- Goal, as before, is to solve $\mathbf{P}'$.

6 Sixth Exercise Set

6.1 Kalman Filter for Geometrical Optics

Ray of light with an initial covariance matrix

$$\mathbf{P} = \begin{pmatrix} \sigma_y^2 & 0 \\ 0 & \sigma_\theta^2 \end{pmatrix}$$

is incident on a convex boundary. What is the covariance matrix after the ray passes through the convex boundary?

- This is again a so-called discrete problem.
- We will use the paraxial apporximation in which $\tan\phi \approx \phi$.

  Encode a ray with the parameters y and θ, packed into the vector $\boldsymbol{y} = (y, \theta)$.

  We write the passage of light through the boundary using a transfer matrix $\boldsymbol{\Phi}$.

$$\begin{pmatrix} y' \\ \theta' \end{pmatrix} = \boldsymbol{\Phi} \begin{pmatrix} y \\ \theta \end{pmatrix}$$

- Without proof, the transfer matrix for a convex boundary reads

$$\mathbf{M} = \begin{pmatrix} 1 & 0 \\ \frac{n_1 - n_2}{n_2 R} & \frac{n_1}{n_2} \end{pmatrix}$$

- Cool: note that $\boldsymbol{y}' = \boldsymbol{\Phi}\boldsymbol{y}$ has the same form as the Kalman dynamics equation $\boldsymbol{y}' = \boldsymbol{\Phi}\boldsymbol{y}$. If we know the optical transfer matrix $\boldsymbol{\Phi}$, then we know the dyanmics matrix $\boldsymbol{\Phi}$.
- In our case the Kalman dynamics matrix is just the optical transfer matrix

$$\boldsymbol{\Phi} = \begin{pmatrix} 1 & 0 \\ \frac{n_1 - n_2}{n_2 R} & \frac{n_1}{n_2} \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ a & b \end{pmatrix}$$

- We aim to find the covariance matrix

$$\mathbf{P'} = \Phi \mathbf{P} \Phi^\top$$

  This product is

$$\begin{aligned} \mathbf{P}' &= \begin{pmatrix} 1 & 0 \\ a & b \end{pmatrix} \begin{pmatrix} \sigma_y^2 & 0 \\ 0 & \sigma_\theta^2 \end{pmatrix} \begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ a & b \end{pmatrix} \begin{pmatrix} \sigma_y^2 & a\sigma_y^2 \\ 0 & b\sigma_\theta^2 \end{pmatrix} \\ &= \begin{pmatrix} \sigma_y^2 & a\sigma_y^2 \\ a\sigma_y^2 & a^2\sigma_y^2 + b^2\sigma_\theta^2 \end{pmatrix} \end{aligned}$$

  Substitute in values of parameters a and b and get

$$\mathbf{P}' = \begin{pmatrix} \sigma_y^2 & \frac{1}{R}\left(\frac{n_1}{n_2} - 1\right)\sigma_y^2 \\ \frac{1}{R}\left(\frac{n_1}{n_2} - 1\right)\sigma_y^2 & \frac{1}{R^2}\left(\frac{n_1}{n_2} - 1\right)^2 \sigma_y^2 + \left(\frac{n_1}{n_2}\right)^2 \sigma_\theta^2 \end{pmatrix}$$

  This is the desired covariance matrix after passage of light through the curved boundary.

- Next consider variances:
$$(\sigma_y^2)' = \sigma_y^2$$
There is no change because there is no change in y for the passage through the (assumed to be infinitely thin) boundary.

  Meanwhile, a considerable change occurs for
$$(\sigma_\theta^2)' = \frac{1}{R^2}\left(\frac{n_1}{n_2} - 1\right)^2 \sigma_y^2 + \left(\frac{n_1}{n_2}\right)^2 \sigma_\theta^2$$
This change is considerable because the angle changes when passinge through the boundary. Note that if $n_1 = n_2$, in which case the angle θ is unchanged, then $(\sigma_\theta^2)' = \sigma_\theta^2$.

6.2 Theory: Continuous Picture

- Dynamics equation is
$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \mathbf{A}\boldsymbol{x} + \boldsymbol{c} + \boldsymbol{\Gamma}\boldsymbol{w}$$
- The filter reads
$$\begin{aligned}\frac{\mathrm{d}\widehat{\boldsymbol{x}}}{\mathrm{d}t} &= \mathbf{A}\widehat{\boldsymbol{x}} + \boldsymbol{c} + \mathbf{K}(\boldsymbol{z} - \mathbf{H}\widehat{\boldsymbol{x}}) \\ \mathbf{K} &= \mathbf{P}\mathbf{H}^\top\mathbf{R}^{-1} \\ \frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} &= \mathbf{A}\mathbf{P} + \mathbf{P}\mathbf{A}^\top + \boldsymbol{\Gamma}\mathbf{Q}\boldsymbol{\Gamma}^\top - \mathbf{P}\mathbf{H}^\top\mathbf{R}^{-1}\mathbf{H}\mathbf{P}\end{aligned}$$
Terms in order: first two terms are effect of dynamics, dynamics noise, and fourth sharpening/optimization because of measurements. We will neglect the fourth term because in our problems we will not consider measurements, only dynamics.

6.3 A Ping-Pong Ball on a Rough Table

Consider a circular table of radius $R = 2\,\mathrm{m}$ with a rough, wavy surface, so that a ball dropped on the table will bounce around in random directions. From experiment, we find that a ping-pong ball dropped vertically onto the table's center with no uncertainty in the radial position (and without any initial velocity in the radial direction) reaches the table's edge in an average time of $5\,\mathrm{s}$. How long will such a ping-pong ball take to reach the table's edge if the ball is dropped from an initial position with an uncertainty $\sigma_{r_0} = 20\,\mathrm{cm}$ in the radial direction?

- When the ball is dropped with zero radial uncertainty, we are given the uncertainty $\sigma_{r_0} = 0\,\mathrm{cm}$ and average time $\tau_0 = 5\,\mathrm{s}$ to reach the edge.

  We aim to to find the time τ_1 to reach the edge when the initial radial uncertainty is $\sigma_{r_1}(0) = 20\,\mathrm{cm}$.
- We will model the ball's random bounces on the rough table as dynamic noise, and analyze the ball's motion with the Kalman filter equation
$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \mathbf{A}\boldsymbol{x} + \boldsymbol{c} + \boldsymbol{\Gamma}\boldsymbol{w}.$$

In the radial direction, the ball's dynamics are given by

$$m\ddot{r} = F(t) \implies \ddot{r} = \frac{F(t)}{m} \equiv w(t),$$

where $F(t)$ represents random forces during bounces. This equation is a second-order equation, which we must convert to a system of two first-order equations in order to apply the Kalman filter. To do this, we define the new variable

$$v \equiv \dot{r} \qquad \implies \qquad \dot{v} = \frac{F(t)}{m}$$

Our vector quantity for the Kalman filter is

$$\boldsymbol{x} = \begin{pmatrix} r \\ v \end{pmatrix}$$

- We aim to write to ball's dynamics in the matrix form

$$\begin{pmatrix} \dot{r} \\ \dot{v} \end{pmatrix} = \mathbf{A} \begin{pmatrix} r \\ v \end{pmatrix} + \boldsymbol{c} + \boldsymbol{\Gamma w}$$

To do this, we write $\dot{r}$ and $\dot{v}$ in terms of r and v in the form

$$\dot{r} = 0 \cdot r + 1 \cdot v + 0$$
$$\dot{v} = 0 \cdot r + 0 \cdot v + w(t),$$

from which we can read off the corresponding matrix equation

$$\begin{pmatrix} \dot{r} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} r \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ w \end{pmatrix},$$

- The vector $\boldsymbol{x}$'s covariance matrix is then given by the Kalman filter equation

$$\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} = \mathbf{AP} + \mathbf{PA}^\top + \boldsymbol{\Gamma Q \Gamma},$$

where we have dropped the measurement term. We proceed in steps, and first compute $\mathbf{AP}$ which reads

$$\mathbf{AP} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} P_{11}(t) & P_{12}(t) \\ P_{21}(t) & P_{22}(t) \end{pmatrix} = \begin{pmatrix} P_{21} & P_{22} \\ 0 & 0 \end{pmatrix}.$$

Note that the components of $\mathbf{P}$ are currently unknown—that's okay, we'll find them later. Next, taking advantage of $\mathbf{P}$'s symmetry (i.e. $\mathbf{P} = \mathbf{P}^\top$) and the general matrix identity $\mathbf{A}^\top \mathbf{B}^\top = (\mathbf{BA})^\top$, we find $\mathbf{PA}^\top$ according to

$$\mathbf{PA}^\top = \mathbf{P}^\top \mathbf{A}^\top = (\mathbf{AP})^\top = \begin{pmatrix} P_{21} & 0 \\ P_{22} & 0 \end{pmatrix}.$$

Finally, from lecture, the dynamic noise term $\boldsymbol{\Gamma Q \Gamma}^\top$ is found from the equation

$$\left\langle \boldsymbol{\Gamma w} \cdot (\boldsymbol{\Gamma w})^\top \right\rangle = \left\langle \boldsymbol{\Gamma w w}^\top \boldsymbol{\Gamma}^\top \right\rangle = \boldsymbol{\Gamma} \left\langle \boldsymbol{w w}^\top \right\rangle \boldsymbol{\Gamma}^\top = \boldsymbol{\Gamma Q \Gamma}^\top.$$

In our case, using $\boldsymbol{\Gamma w} = (0, w)^\top$, this above equation reads

$$\boldsymbol{\Gamma Q \Gamma}^\top = \left\langle \begin{pmatrix} 0 \\ w \end{pmatrix} \cdot \begin{pmatrix} 0 & w \end{pmatrix} \right\rangle = \begin{pmatrix} 0 & 0 \\ 0 & \langle w^2 \rangle \end{pmatrix} \equiv \begin{pmatrix} 0 & 0 \\ 0 & Q \end{pmatrix},$$

where we have defined $Q \equiv \langle w^2 \rangle$ to denote the variance of dynamic noise.

Covariance Matrix

- We can write the equation for $\mathbf{P}$ now. This is

$$\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} = \begin{pmatrix} P_{21} & P_{22} \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} P_{21} & 0 \\ P_{22} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & Q \end{pmatrix} = \begin{pmatrix} 2P_{21} & P_{22} \\ P_{22} & Q \end{pmatrix}.$$

  So our goal has been to find differential equation for $\mathbf{P}$, so that we can solve for the elements of $\mathbf{P}$. Basically, we have

$$\begin{pmatrix} \dot{P}_{11} & \dot{P}_{12} \\ \dot{P}_{21} & \dot{P}_{22} \end{pmatrix} = \begin{pmatrix} 2P_{21} & P_{22} \\ P_{22} & Q \end{pmatrix}$$

- The above matrix equation correponds to the three differential equations

$$\dot{P}_{11} = 2P_{21} = 2P_{12} \qquad \dot{P}_{12} = P_{22} \qquad \dot{P}_{22} = Q,$$

  with the general solution

$$\begin{aligned} P_{22}(t) &= Qt + P_{22}(0) \\ P_{12}(t) &= \frac{1}{2}Qt^2 + P_{22}(0) \cdot t + P_{12}(0) \\ P_{11}(t) &= \frac{1}{3}Qt^3 + P_{22}(0)t^2 + 2P_{12}(0) \cdot t + P_{11}(0), \end{aligned} \tag{6.1}$$

  where $P_{11}(0)$, $P_{12}(0)$ and $P_{22}(0)$ correspond to the initial variance in radial position, initial covariance $\sigma_{rv}(0)$ and initial variance in radial velocity, respectively.

Finding Variance of Dynamic Noise Q

- We first find Q using the first set of initial conditions, in which $\sigma_{r_0}^2 = 0\,\mathrm{cm}$ and the ball reaches the table's edge at $t = \tau_0$. In this case all initial variables are known exactly (without uncertainty), and, writing all steps out, we have

$$\mathbf{P}(0) = \begin{pmatrix} P_{11}(0) & P_{12}(0) \\ P_{12}(0) & P_{22}(0) \end{pmatrix} = \begin{pmatrix} \sigma_{r_0}^2 & \sigma_{r_0 v_0} \\ \sigma_{r_0 v_0} & \sigma_{v_0}^2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

  In this case Equation 6.1 for $P_{11}(t)$ reduces to

$$P_{11}(t) = \frac{1}{3}Qt^3. \tag{6.2}$$

- Interpretation: the randomly bouncing ball reaching the table's edge $r = R$ at time τ_0 really mean's the uncertainty σ_r in the ball's radial position at $t = \tau_0$ is $\sigma_r(\tau_0) = R$. Combining this consideration with Equation 6.2 produces

$$P_{11}(\tau_0) = \sigma_r^2(\tau_0) = \frac{1}{3}Q\tau_0^3 \implies Q = \frac{3R^2}{\tau_0^3} \tag{6.3}$$

Last Part

- From the problem's instructions, we aim to find the time τ_1 at which the ball reaches the table's edge when the initial radial uncertainty is $\sigma_r(0) = \sigma_{r_0}$. The corresponding covariance matrix is

$$\mathbf{P}(t=0) = \begin{pmatrix} \sigma_{r_0}^2 & 0 \\ 0 & 0 \end{pmatrix}.$$

  In this case, Equation 6.1 for the variance in radial position reads

$$\begin{aligned} P_{11}(t) &= \frac{1}{3}Qt^3 + P_{11}(0) = \frac{1}{3}\left(\frac{3R^2}{\tau_0^3}\right)t^3 + \sigma_{r_0}^2 \\ &= R^2\left(\frac{t}{\tau_0}\right)^3 + \sigma_{r_0}^2, \end{aligned}$$

  where we have substituted in Q from Equation 6.3.

- Again interpreting R as the uncertainty $\sigma_r(t)$ in radial position when the ball reaches the table's edge, this time at $t = \tau_1 \implies \sigma_r(\tau_1) = R$, we get the equation

$$P_{11}(\tau_1) = \sigma_r^2(\tau_1) = R^2 = R^2\left(\frac{\tau_1}{\tau_0}\right)^3 + \sigma_{r_0}^2,$$

  which we then solve for τ_1 to get

$$\tau_1 = \tau_0\left(\frac{R^2 - \sigma_{r_0}^2}{R^2}\right)^{1/3} = \tau_0\left(1 - \frac{\sigma_{r_0}^2}{R^2}\right)^{1/3}.$$

  Note that, if $\sigma_{r_0} = 0$ (no unceratinty in initial radial position) we recover the earlier result $\tau_1 = \tau_0$. Meanwhile, if $\sigma_{r_0} \neq 0$ we have $\tau_1 < \tau_0$, in which case the ball reaches the table's edge sooner. This should make sense, because the ball's initial position is statistically more spread out from the table's center, and so has to travel less distance to reach the edge.

  TODO: notation for subscritps on σ_{r_0} which is used twice.

- For our concrete data, with $R = 2\,\mathrm{m}$ and $\sigma_r(0) = 20\,\mathrm{cm}$ and $\tau_0 = 5\,\mathrm{s}$, we have

$$\tau = 5\,\mathrm{s} \cdot \left(1 - \frac{(20\,\mathrm{cm})^2}{(2\,\mathrm{m})^2}\right)^{1/3} \approx 4.93\,\mathrm{s}.$$

7 Seventh Exercise Set

7.1 Ball Falling Through an Optical Gate

Two circular optical gates, both of radius R, are placed one above the other and separated by a vertical distance h. We perform an experiment in which we drop balls from a height h above the upper gate, and observe the roughly two thirds of all balls fall through the upper gate. Estimate the fraction of ball that also fall through the lower get. Assume the balls' initial position and velocity are exactly known.

- Define transverse direction as r. Define t_1 time at gate 1 and t_2 time at gate 2.
- Next interpret random forces in radial direction: $\langle F(t)\rangle = 0\,\mathrm{N}$. We then wrote $\langle r(t)\rangle = 0$—interpret. Ah makes sense—r is transverse position; zero transverse force should be zero transverse displacement.
- Initial (transverse) position and speed of ball are exactly known: interpretation is $\mathbf{P}(0) = \mathbf{0}$.
- We assume $r \sim \mathcal{N}(0, \sigma_r^2(t))$. Note that the variance increases, and the corresponding Gaussian widens with time. Because $x \sim \mathcal{N}$, we can find the probability P_2 with standard Gaussian CDF.

  First define the standardized variable

  $$\rho \equiv \frac{r - \langle r\rangle}{\sigma_r} = \frac{r-0}{\sigma_r} = \frac{r}{\sigma_r}.$$

  We then have

  $$P_2 = P(-R \le r(t_2) \le R) = \Phi(u(R)) - \Phi(u(-R)) = \Phi\left(\frac{R}{\sigma_r(t_2)}\right) - \Phi\left(\frac{-R}{\sigma_r(t_2)}\right)$$
  $$= 2\Phi\left(\frac{R}{\sigma_r(t_2)}\right) - 1.$$

Kalman Filter Dynamics

- Recall Kalman filter:
  $$\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} = \mathbf{A}\mathbf{P} + \mathbf{P}\mathbf{A}^\top + \mathbf{\Gamma}\mathbf{Q}\mathbf{\Gamma}^\top$$
  We will consider the transverse direction only, since we aim to find uncertainty in radial position r. Start
  $$m\ddot{r} = F(t) \implies \ddot{r} = \frac{F(t)}{m} \equiv w$$
  Write random forces as random noise. Having written random noise, introduce Kalman notation:
  $$\dot{\boldsymbol{r}} = \begin{pmatrix}\dot{r}\\ \dot{v}\end{pmatrix} = \mathbf{A}\boldsymbol{r} + \boldsymbol{c} + \mathbf{\Gamma}\boldsymbol{w}$$
  Used Kalman dynamic equation. Reusing previous problem
  $$\begin{pmatrix}\dot{r}\\ \dot{v}\end{pmatrix} = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix}r\\ v\end{pmatrix} + \begin{pmatrix}0\\ 0\end{pmatrix} + \begin{pmatrix}0\\ w\end{pmatrix}$$

- The result is

$$\dot{\mathbf{P}} = \mathbf{AP} + \mathbf{PA}^\top + \mathbf{\Gamma Q \Gamma}^\top = \begin{pmatrix} P_{21} & P_{22} \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} P_{21} & 0 \\ P_{22} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & Q \end{pmatrix}$$
$$= \begin{pmatrix} 2P_{21} & P_{22} \\ P_{22} & Q \end{pmatrix}$$

  The coresponding system of equations is

  The general solution is

$$\begin{aligned} \dot{P}_{22} = Q \quad &\implies P(t) = Qt + P_{22}(0) \\ \dot{P}_{12} = P_{22} \quad &\implies P_{12}(t) = \frac{1}{2}Qt^2 + P_{22}(0) \cdot t + P_{12}(0) \\ \dot{P}_{11} = 2P_{21} \quad &\implies P_{11}(t) = \frac{1}{3}Qt^3 + P_{22}(0) \cdot t^2 + 2P_{12}(0) \cdot t + P_{11}(0). \end{aligned}$$

  From the initial condition $\mathbf{P}(0) = \mathbf{0}$ (no initial uncertainty in any variables), the solution simplifies to

$$P_{22}(t) = Qt \qquad P_{12} = \frac{1}{2}Qt^2 \qquad P_{11}(t) = \frac{1}{3}Qt^3.$$

Finding Variance of Dynamic Noise Q

- To find Q, we recall the given experimental data

$$\mathcal{P}_1 = \mathcal{P}(-R \leq r(t_1) \leq R) \approx 2/3.$$

  In other words, two third of all balls fall through first gate. The above probability corresponds to a Gaussian 1-σ interval, from which we can conclude

$$1 \cdot \sigma_r(t_1) \approx R.$$

- From $\sigma_r(t_1) = R$ and $P_{11}(t) = \sigma_r(t)^2$ we can then conclude

$$P_{11}(t_1) \equiv \sigma_r^2(t_1) \approx R^2 = \frac{1}{3}Qt_1^3 \implies Q \approx \frac{3R^2}{t_1^3}$$

Last Part

- Goal is to find $\sigma_r(t_2)$. From known Q we have

$$P_{11}(t_2) = \sigma_r^2(t_2) = \frac{1}{3}Qt_2^3 = \frac{1}{3}\left(\frac{3R^2}{t_1^3}\right)t_2^3 = R^2\left(\frac{t_2}{t_1}\right)^3.$$

  We find the times t_1 and t_2 from known free fall kinematics, which give

$$t_1 = \sqrt{\frac{2h}{g}} \qquad \text{and} \qquad t_2 = \sqrt{\frac{2 \cdot (2h)}{g}} \implies \frac{t_2}{t_1} = \sqrt{2}.$$

- With known times, transverse position variance is then

$$\sigma_r^2(t_2) = R^2 \left(\frac{t_2}{t_1}\right)^3 = 2^{3/2} \cdot R^2.$$

The problem asks for uncertainty, which is

$$\sigma_r(t_2) = \sqrt{\sigma_r^2(t_2)} = 2^{3/4} \cdot R.$$

This is uncertainty in transverse position of falling ball when they reach the second gate. From the standard Gaussian distribution we then have

$$\mathcal{P}(-R \leq r(t_2) \leq R) = 2\Phi\left(\frac{R}{\sigma_r(t_2)}\right) - 1 = 2 \cdot F(2^{-3/4}) - 1 \approx 2 \cdot F(0.595) - 1 \approx 0.45$$

7.2 Motion in Viscous Fluid

A small ball of mass m falls slowly in a viscous fluid; the ball's dynamics in the vertical direction are given by $F/m = -\beta v + w(t)$, where $-\beta v$ is a velocity-dependent damping term, while the term $w(t)$ represents dynamic noises. Use Kalman filter with vector $\boldsymbol{x} = (z, v)^\top$ to follow the ball's motion through fluid. The initial and large-time uncertainties in velocity are known to be

$$\sigma_v^2(0) \equiv \sigma_{v_0}^2 \qquad \text{and} \qquad \lim_{t\to\infty} \sigma_v^2(t) = \frac{1}{4}\sigma_{v_0}^2.$$

Determine the uncertainty in the ball's velocity at the time t_0 when $\beta t_0 = 1$.

- We choose our coordinate system so the z axis points in the direction of gravitational acceleration $\boldsymbol{g}$.

  After accounting for weight and buoyancy, which we write in terms of an effective acceleration g_{eff}, the ball's full dynamics in the z direction read

$$m\ddot{z} = mg_{\text{eff}} - m\beta v + F(t),$$

  where $F(t)$ is random noise in the vertical direction. We then solve for $\ddot{z}$ to get

$$\ddot{z} = g_{\text{eff}} - \beta v + \frac{F(t)}{m} \equiv g_{\text{eff}} - \beta v + w(t), \tag{7.1}$$

  where we have defined the dynamic noise term $w \equiv F/m$.

- Our next step is to write the dynamics in the Kalman form $\dot{\boldsymbol{x}} = \mathbf{A}\boldsymbol{x} + \boldsymbol{c} + \boldsymbol{\Gamma w}$. To do this, we decompose $\ddot{z}$ into two variables

$$\dot{z} = v \qquad \text{and} \qquad \dot{v} = g_{\text{eff}} - \beta v + w,$$

  in terms of which, in matrix form, Equation 7.1 becomes

$$\begin{pmatrix} \dot{z} \\ \dot{v} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ 0 & -\beta \end{pmatrix}}_{\mathbf{A}} \begin{pmatrix} z \\ v \end{pmatrix} + \underbrace{\begin{pmatrix} 0 \\ g_{\text{eff}} \end{pmatrix}}_{\boldsymbol{c}} + \underbrace{\begin{pmatrix} 0 \\ w(t) \end{pmatrix}}_{\boldsymbol{\Gamma w}}.$$

- We will then find the covariance matrix $\mathbf{P}(t)$ from the differential equation

$$\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} = \mathbf{AP} + \mathbf{PA} + \mathbf{\Gamma Q \Gamma}^\top. \tag{7.2}$$

We first compute the product $\mathbf{AP}$ according to

$$\mathbf{AP} = \begin{pmatrix} 0 & 1 \\ 0 & -\beta \end{pmatrix} \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} = \begin{pmatrix} P_{21} & P_{22} \\ -\beta P_{21} & -\beta P_{22} \end{pmatrix}.$$

Next, using $\mathbf{P} = \mathbf{P}^\top$, we find $\mathbf{PA}$ from

$$\mathbf{PA}^\top = \mathbf{P}^\top \mathbf{A}^\top = (\mathbf{AP})^\top = \begin{pmatrix} P_{21} & -\beta P_{21} \\ P_{22} & -\beta P_{22} \end{pmatrix}.$$

Finally, we compute

$$\mathbf{\Gamma Q \Gamma}^\top \equiv \left\langle (\mathbf{\Gamma} \boldsymbol{w}) \cdot (\mathbf{\Gamma} \boldsymbol{w})^\top \right\rangle = \begin{pmatrix} 0 & 0 \\ 0 & \langle w^2(t) \rangle \end{pmatrix} \equiv \begin{pmatrix} 0 & 0 \\ 0 & Q \end{pmatrix},$$

where we have defined $Q \equiv \langle w^2 \rangle$ as the variance of dynamic noise. After adding the terms in Equation 7.2, the complete differential equation for $\mathbf{P}$ is then

$$\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} = \begin{pmatrix} \dot{P}_{11} & \dot{P}_{12} \\ \dot{P}_{21} & \dot{P}_{22} \end{pmatrix} \begin{pmatrix} 2P_{21} & P_{22} - \beta P_{21} \\ P_{22} - \beta P_{21} & Q - 2\beta P_{22} \end{pmatrix}$$

- Recall we aim to find the velocity variance $\sigma_v^2(t)$, which corresponds to the covariance term $P_{22}(t)$. The relevant differential equation is

$$\dot{P}_{22} \equiv \frac{\mathrm{d}P_{22}}{\mathrm{d}t} = Q - 2\beta P_{22}. \tag{7.3}$$

We solve the equation by defining the new variable

$$u \equiv Q - 2\beta P_{22} \implies \frac{\mathrm{d}u}{\mathrm{d}t} = -2\beta \frac{\mathrm{d}P_{22}}{\mathrm{d}t}$$

In terms of u, Equation 7.3 reads

$$\frac{\mathrm{d}P_{22}}{\mathrm{d}t} = Q - 2\beta P_{22} \iff -\frac{1}{2\beta}\frac{\mathrm{d}u}{\mathrm{d}t} = u \implies \frac{\mathrm{d}u}{u} = -2\beta\, \mathrm{d}t.$$

The final equation has the familiar exponential solution

$$u(t) = Ce^{-2\beta t},$$

where C is a constant. The corresponding solution for P_{22}, using $u = Q - 2\beta P_{22}$, is

$$Q - 2\beta P_{22}(t) = Ce^{-2\beta t} \tag{7.4}$$

- We find the constant C from the initial condition $P_{22}(0) = \sigma_{v_0}^2$, which gives

$$Q - 2\beta P_{22}(0) = C \cdot 1 \implies C = Q - 2\beta\sigma_{v_0}^2.$$

We then substitute C into Equation 7.4 to get

$$Q - 2\beta P_{22}(t) = \left(Q - 2\beta\sigma_{v_0}^2\right)e^{-2\beta t} \tag{7.5}$$

We then find the variance of dynami noise Q from the condition $\sigma_v^2 = P_{22}(t) \to \frac{1}{4}\sigma_{v_0}^2$ as $t \to \infty$, which gives

$$Q - 2\beta P_{22}(\infty) = 0 \implies Q = 2\beta P_{22}(\infty) = \frac{\beta\sigma_{v_0}^2}{2}.$$

We then substitute Q into Equation 7.5, which then reads

$$\frac{\beta\sigma_{v_0}^2}{2} - 2\beta P_{22}(t) = \left(\frac{\beta\sigma_{v_0}^2}{2} - 2\beta\sigma_{v_0}^2\right) e^{-2\beta t}.$$

After dividing through by 2β and rearranging, this gives

$$P_{22}(t) = \frac{\sigma_{v_0}^2}{4} - \left(\frac{\sigma_{v_0}^2}{4} - \sigma_{v_0}^2\right) e^{-2\beta t} = \frac{\sigma_{v_0}^2}{4}\left(1 + 3e^{-2\beta t}\right).$$

Finally, we aim to find $P_{22}(t_0)$, when $\beta t_0 = 1$, in which case we have

$$P_{22}(t_0) = \sigma_v^2(t_0) = \frac{\sigma_{v_0}^2}{4}(1 + 3e^{-2\beta t_0}) = \frac{\sigma_{v_0}^2}{4}(1 + 3e^{-2}).$$

The corresponding uncertainty σ_v in the ball's velocity is

$$\sigma_v(t_0) = \frac{\sigma_{v_0}}{2}\left(1 + 3e^{-2}\right)^{1/2}.$$

8 Eighth Exercise Set

8.1 Theory: Introduction to Sensors

- $x(t)$ is output of sensor. Input is $z(t)$.

  A differential equation relates the input to the outut.

- It is more convenient to solve these differential equations with the Laplace transformation, which transforms differential equations in the time domain to algebraic equations in the Laplace domain s.

- Schematic: the time-domain procedure

  $$z(t) \to \text{differential equation} \to x(t),$$

  where the input $z(t)$ is mapped to the output $x(t)$ via a differential equation, becomes

  $$z(s) \cdot H(s) = X(s)$$

  in the Laplace domain, where $H(s)$ is the sensor's transfer function.

- The Laplace transformation of a time-domain function $f(t)$ is defined by

  $$\mathcal{L}(f(t)) = \lim_{\epsilon \to 0^-} \int_{\epsilon}^{\infty} f(t)e^{-st}\,\mathrm{d}t \equiv \int_{0^-}^{\infty} f(t)e^{-st}\,\mathrm{d}t = F(s).$$

  By convention, we will write time domain functions in lower case (e.g. $f(t)$) and their Laplace transforms in uppercase (e.g. $F(s)$).

- Following is an important identity, involving the Laplace transformation of a function's time derivative. Using integration by parts, we have

  $$\begin{aligned}\mathcal{L}(\dot{f}(t)) &= \int_{0^-}^{\infty} \dot{f}e^{-st}\,\mathrm{d}t = f(t)e^{-st}\Big|_{0^-}^{\infty} - \int_{0^-}^{\infty} (-s \cdot f(t)e^{-st})\,\mathrm{d}t \\ &= -f(0^-) + s\int_{0^-}^{\infty} f(t)e^{-st}\,\mathrm{d}t \\ &= -f(0^-) + sF(s).\end{aligned}$$

  We have assumed $t(t)$ decays rapidly enough at infinity that $f(t)e^{-st} \to 0$ as $t \to \infty$.

  In practice, we are interested in a sensor's response only from some given point in time forward. In this case, we can choose **TODO:** what is f, sensor input or output? the origin so that $f(0) = 0$, in which case

  $$\mathcal{L}(\dot{f}(t)) = s \cdot F(s) \qquad \text{(this assumes } f(0) = 0\text{)}.$$

  This identity generalizes straightforwardly to an n-th order time derivative as

  $$\mathcal{L}\left(f^{(n)}(t)\right) = s^n F(s).$$

- In practice use tables to find Laplace transformations of common functions. More so, find inverse transformations by comparing a function of s to its known inverse or whatever blah phrase better.

$f(t) = \mathcal{L}^{-1}(F(s))$	$F(s) = \mathcal{L}(f(t))$	$f(t) = \mathcal{L}^{-1}(F(s))$	$F(s) = \mathcal{L}(f(t))$
e^{at}	$\dfrac{1}{s-a}$	1	$1/s$
t^n	$\dfrac{n!}{s^{n+1}}$	$\delta(t)$	1
$\sin \omega t$	$\dfrac{\omega}{s^2+\omega^2}$	$\cos \omega t$	$\dfrac{s}{s^2+\omega^2}$
$\Theta(t-t_0)$	$\dfrac{1}{s}e^{-t_0 s}$	$f(t-t_0)\Theta(t-t_0)$	$F(s)e^{-t_0 s}$
$g(t)e^{at}$	$G(s-a)$	$\frac{\mathrm{d}^n}{\mathrm{d}t^n}g(t)$	$s^n G(s)$
$-t \cdot g(t)$	$\frac{\mathrm{d}}{\mathrm{d}s}G(s)$		

Table 2: Common Laplace transforms. The symbol $\Theta(t)$ denotes the Heaviside step function.

First Order Sensors

- All first-order sensors obey a differential equation of the form

  $$z(t) = \tau \dot{x}(t) + x(t),$$

  where τ is called the sensor's *time constant*. To find the sensor's output $x(t)$, we first take Laplace transformation of above equation and get

  $$\tau s X(s) + X(s) = Z(s) \implies X(s) = \frac{Z(s)}{1+\tau s},$$

  which motivates the definition of a first order sensor's *transfer function* as

  $$H(s) = \frac{1}{1+\tau s}.$$

  We then find the sensor's output $x(t)$ with the inverse Laplace transformation

  $$x(t) = \mathcal{L}^{-1}\big\{H(s) \cdot Z(s)\big\}.$$

8.2 Response of First Order Sensors to Common Signals

8.2.1 Response to a Delta Function

Find the response of a first-order sensor to a delta function input $z(t) = \delta(t)$.

- Begin with Laplace transform of general expression.

  $$X(s) = H(s)Z(s) = \frac{1}{1+\tau s} \cdot Z(s).$$

  From Table **??**, the input signal's Laplace transform is

  $$Z(s) = \mathcal{L}(z(t)) = \mathcal{L}(\delta(t)) = 1 \implies X(s) = H(s) \cdot 1 = \frac{1}{1+\tau s}.$$

- We find the time-domain sensor output $x(t)$ with the inverse transform

$$x(t) = \mathcal{L}^{-1}\left\{\frac{1}{1+\tau s}\right\}, \tag{8.1}$$

where we find the inverse by using the Table **??** identity

$$\mathcal{L}\{e^{at}\} = \frac{1}{s-a} \implies \mathcal{L}^{-1}\left\{\frac{1}{s-a}\right\} = e^{at}. \tag{8.2}$$

To apply this identity to Equation 8.1, we rewrite

$$\frac{1}{1+\tau s} = \frac{1}{\tau}\cdot\frac{1}{s+1/\tau},$$

which matches Equation 8.2 if we choose $x = -1/\tau$. The sensor's output $x(t)$ is thus

$$x(t) = \mathcal{L}^{-1}\left\{\frac{1}{1+\tau s}\right\} = \frac{1}{\tau}e^{-t/\tau}.$$

8.2.2 Response to Linearly Increasing Input

Find the response of a first-order sensor to the linearly increasing input $z(t) = kt$.

- From Table 8.2, the input signal's Laplace transform is

$$Z(s) = k\mathcal{L}(t) = k\cdot\frac{1}{s^2} \implies X(s) = H(s)Z(s) = \frac{1}{1+\tau s}\frac{k}{s^2}$$

- To find $x(t)$, we first need to transform $X(s)$ into an easily-invertible form. We do this with the partial fraction decomposition.

$$X(s) = \frac{1}{1+\tau s}\frac{k}{s^2} = \frac{A}{1+\tau s} + \frac{Bs+C}{s^2}. \tag{8.3}$$

Forming a common denominator leads to the equality

$$\begin{aligned} k &= As^2 + (1+\tau s)(Bs+C) = As^2 + \tau Bs^2 + Bs + \tau Cs + C \\ &= (A+\tau B)s^2 + s(B+\tau C) + C, \end{aligned}$$

and equating the powers of s leads to the system of equations

$$\underbrace{A+\tau B = 0}_{\text{from } s^2} \qquad \underbrace{B+\tau C = 0}_{\text{from } s^1} \qquad \underbrace{C = k}_{\text{from } s^0}.$$

This system has the solution $A = \tau^2 k$, $B = -\tau k$ and $C = k$. The decomposition of $X(s)$ in Equation 8.3 is then

$$X(s) = \frac{\tau^2 k}{1+\tau s} - \frac{\tau k}{s} + \frac{k}{s^2} = \frac{k}{\tau}\frac{1}{s+1/\tau} - \tau k\cdot\frac{1}{s} + k\cdot\frac{1}{s^2}.$$

From Table **??**, the inverse transformation is

$$\begin{aligned} x(t) &= \mathcal{L}^{-1}(X(s)) = k\tau\mathcal{L}^{-1}\left\{\frac{1}{s+1/\tau}\right\} - \tau k\mathcal{L}^{-1}\left\{\frac{1}{s}\right\} + k\mathcal{L}^{-1}\left\{\frac{1}{s^2}\right\} \\ &= k\tau e^{-t/\tau} - \tau k + kt \\ &= kt - k\tau(1-e^{-t/\tau}). \end{aligned}$$

For large times, i.e. $t \gg \tau$, the sensor's output reads

$$\lim_{t \gg \tau} x(t) = kt - k\tau$$

In other words, the first-order sensor's output correctly captures the input's linear functional form $z(t) \propto kt$, but adds a constant offset $-k\tau$.

8.2.3 Response to Quadratically Increasing Input

Find the response of a first-order sensor to the quadratically increasing input $z(t) = at^2$.

- From Table **??**, the input's Laplace transform is

$$Z(s) = \mathcal{L}\{at^2\} = \frac{2a}{s^3} \implies X(s) = H(s) \cdot Z(s) = \frac{1}{1+\tau s} \cdot \frac{2a}{s^3}$$

As in the previous problem, we simply $X(s)$ with partial fraction decomposition, giving

$$X(s) = \frac{1}{1+\tau s} \cdot \frac{2a}{s^3} = \frac{A}{1+\tau s} + \frac{Bs^2 + Cs + D}{s^3}.$$

We then form common denominator, producing the equality

$$2a = (A + \tau B)s^3 + (B + \tau C)s^2 + (C + D\tau)s + D,$$

before equating coefficients s to get the system of equations

$$A + \tau B = 0 \qquad B + \tau C = 0 \qquad C + \tau D = 0 \qquad 2a = D.$$

The solutions are

$$D = 2a \qquad C = -2a\tau \qquad B = 2a\tau^2 \qquad A = -2a\tau^3,$$

and the corresponding decomposition for $X(s)$ is

$$X(s) = \frac{2a}{s^3} - \frac{2a\tau}{s^2} + \frac{2a\tau^2}{s} - \frac{2a\tau^3}{1+\tau s}.$$

- We then find the time-domain output with the inverse transformation

$$\begin{aligned} x(t) = \mathcal{L}^{-1}(X(s)) &= 2a\mathcal{L}^{-1}\left\{\frac{1}{s^3}\right\} - 2a\tau\mathcal{L}^{-1}\left\{\frac{1}{s^2}\right\} + 2a\tau^2\left\{\frac{1}{s}\right\} - 2a\tau^2 e^{-t/\tau} \\ &= at^2 - 2a\tau t + 2a\tau^2\left(1 - e^{-t/\tau}\right). \end{aligned}$$

Note that for $t \gg \tau$, the sensor's output reads

$$\lim_{t \gg \tau} x(t) = at^2 - 2a\tau t + 2a\tau^2$$

Note that the first-order sensor's output fails to correctly track the input $z(t) = at^2$, even for large times, because of the linearly increasing term $-2a\tau t$.

9 Ninth Exercise Set

9.1 Theory: First-Order Sensor with Non-Zero Initial Input

Input is Non-Zero at $t = 0$

- So far, we had assumed the sensor's input was zero when $t = 0$, i.e. that $z(0) = 0$. We now consider the more general case in which $z(0^-) = z_0 \neq 0$.
- Recall that a first-order sensor obeys the general equation

$$\tau\dot{x} + x = z. \tag{9.1}$$

  If $z(0^-) \neq 0$, we proceed by defining the new variables $\chi = x - x_0$ and $\zeta = z - x_0$, in terms of which Equation 9.1 becomes

$$\tau\dot{\chi} + \chi = \zeta. \tag{9.2}$$

- For example, for a linear input $z(t) = z_0 + xt$, a sensor "sees" the effective input

$$z(t) = (\zeta(t) + x_0) = z_0 + kt \implies \zeta(t) = (z_0 - x_0) + kt$$

  Procedure: with ζ known, solve for $\chi(t)$ using Equation 9.2, then transform back with $x(t) = x_0 + \chi(t)$ for $t \geq 0$.

Signal Begins at $t = t_0 \neq 0$

- Next consideration: what happens if a signal starts at $t = t_0 > 0$ and not at $t = 0$. In this case we write the input in the form

$$z(t) = k(t - t_0) \cdot \Theta(t - t_0), \tag{9.3}$$

  where $\Theta(t)$ is the Heaviside step function. We then apply the general property

$$\mathcal{L}\{f(t - t_0)H(t - t_0)\} = F(s)e^{-t_0 s}$$

  in terms of which the Laplace transformation of Equation 9.3 reads

$$Z(s) = \mathcal{L}\{z(t)\} = \frac{k}{s^2}e^{-t_0 s}.$$

  With $Z(s)$ known, we would then follow the same procedure as in previous problems: we would first find the sensor's response $X(s)$ in the Laplace domain from

$$X(s) = H(s) \cdot Z(s) = \frac{1}{1 + \tau s} \cdot \frac{k}{s^2}e^{-t_0 s},$$

  then find the time-domain response $x(t)$ with an inverse Laplace transform of $X(s)$.

9.2 Example: First-Order Sensor with Non-Zero Initial Input

Find the response of a first-order sensor the linear input $z(t) = z_0 + kt$ where $z_0 \neq 0$.

- Important consideration: the input $z(t) = z_0 + kt$ is nonzero at $t = 0$. As a result, the sensor's output is also nonzero at $t = 0$, i.e. $x(t = 0) \equiv x_0 \neq 0$. We must take $x(0) \neq 0$ into consideration during the Laplace transform of the sensor's differential equation. The sensor's differential equation reads

  $$z(t) = z_0 + kt = \tau\dot{x} + x.$$

  Combining the general identity $\mathcal{L}(\dot{f}) = sF(s) - f(0^-)$ with the result $x(0) = x_0 \neq 0$, the differential equation's Laplace transform is

  $$\tau \cdot \big[sX(s) - x(0)\big] + X(s) = Z(s) \implies X(s) = \frac{Z(s) + \tau x_0}{1 + \tau s}.$$

  The input's transform, using the standard transforms in Table 2, is

  $$Z(s) = \mathcal{L}\big\{z(t)\big\} = \mathcal{L}\{z_0\} + k\mathcal{L}\{t\} = \frac{z_0}{s} + \frac{k}{s^2} = \frac{k + z_0 s}{s^2}.$$

  Before writing down $X(s)$, we first make the auxiliary calculation

  $$Z(s) + \tau x_0 = \frac{k + z_0 s}{s^2} + \tau x_0 = \frac{\tau x_0 s^2 + z_0 \cdot s + k}{s^2},$$

  in terms of which the sensor's Laplace domain output reads

  $$X(s) = \frac{Z(s) + \tau x_0}{1 + \tau s} = \frac{1}{1 + \tau s} \cdot \frac{\tau x_0 s^2 + z_0 s + k}{s^2}.$$

- As usual, we will find $x(t)$ from the partial fraction decomposition of $X(s)$, which reads

  $$X(s) = \frac{A}{1 + \tau s} + \frac{Bs + C}{s^2}$$

  We then form a common denominator and combine like terms to get the equality

  $$\tau x_0 s^2 + z_0 z + k = (A + \tau B)s^2 + (B + \tau C)s + C,$$

  and equate coefficients of powers of s to get the equations

  $$C = k \qquad B + \tau C = z_0 \qquad A + \tau B = \tau x_0$$

  The solutions for A, B and C read

  $$A = \tau x_0 - \tau z_0 + k\tau^2 \qquad B = z_0 - k\tau \qquad C = k,$$

  and the partial fraction decomposition of $X(s)$ reads

  $$X(s) = \frac{k}{s^2} + \frac{z_0 - k\tau}{s} + \frac{x_0 - z_0 + k\tau}{s + 1/\tau}.$$

- The inverse transform, after some technically nonessential rearranging, comes out to

  $$\begin{aligned} x(t) = \mathcal{L}^{-1}\big\{X(s)\big\} &= kt + (z_0 - k\tau) + (x_0 - z_0 + k\tau)e^{-t/\tau} \\ &= kt + x_0 + (z_0 - x_0) - k\tau + (x_0 - z_0)e^{-t/\tau} + k\tau e^{-t/\tau} \\ &= kt + x_0 - (x_0 - z_0)\left(1 - e^{-t/\tau}\right) - k\tau\left(1 - e^{-t/\tau}\right). \end{aligned}$$

  Note that $x(0) = x_0$, as expected.

9.3 Theory: Second-Order Sensors

- A second-order sensor obeys the general differential equation

$$\ddot{x} + 2\zeta\omega_0\dot{x} + \omega_0^2 x = \omega_0^2 z + 2\zeta\omega_0\dot{z},$$

where ω_0 and ζ are the sensor's resonant frequency and damping coefficient, respecitively.

- In practice, sensors tend to be sensitive only to either the input z or the input's derivative $\dot{z}$, but not both. We will assume dependence on only z, in which case the sensor's equation reduces to

$$\ddot{x} + 2\zeta\omega_0\dot{x} + \omega_0^2 x = \omega_0^2 z.$$

- Without proof (see the accompanying lecture notes **TODO:** add link), the sensor is optimally damped when $\zeta = \zeta_{\text{opt}} = \frac{1}{\sqrt{2}}$.

- Assuming $x(0^-) = 0$ and $z(0^-) = 0$, the sensor's equation's Laplace transform reads

$$s^2X(s) + 2\zeta\omega_0 sX(s) + \omega_0^2 X(s) = \omega_0^2 Z(s)$$

We then solve for $X(s)$ to get

$$X(s) = \frac{\omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2} Z(s),$$

which motivates the second-order transfer function as

$$H(s) \equiv \frac{\omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2}.$$

9.4 Example: Response of Second-Order Sensors to Common Inputs

9.4.1 Delta Function Input

Compute the response of a second-order sensor to the delta function input $z(t) = \delta(t)$.

- From Table 2, the input's Laplace transform is

$$Z(s) = \mathcal{L}\{\delta(t)\} = 1.$$

The sensor's output in the Laplace domain is then simply

$$X(s) = H(s) \cdot Z(s) = \frac{\omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2} \cdot 1.$$

- For reasons that will soon be clear, we then write $X(s)$'s denominator in the form

$$\begin{aligned} s^2 + 2\zeta\omega_0 s + \omega_0^2 &= (s + \zeta\omega_0)^2 - \zeta^2\omega_0^2 + \omega_0^2 \\ &= (s + \zeta\omega_0)^2 + \omega_0^2(1 - \zeta^2)^2, \end{aligned} \tag{9.4}$$

in terms of which $X(s)$ reads

$$X(s) = \frac{\omega_0^2}{(s + \zeta\omega_0)^2 + \omega_0^2(1 - \zeta^2)} = \frac{\omega_0}{\sqrt{1 - \zeta^2}} \cdot \frac{\omega_0\sqrt{1 - \zeta^2}}{(s + \zeta\omega_0)^2 + \omega_0^2(1 - \zeta^2)}. \tag{9.5}$$

We then compare this expression for $X(s)$ to identity (constructed from Table 2)

$$\mathcal{L}\left\{e^{at}\sin\omega t\right\} = \frac{\omega}{(s-a)^2+\omega^2},$$

which motivates the choice of $a = -\zeta\omega_0$ and $\omega^2 = \omega_0^2(1-\zeta^2)$, and in terms of which the sensor's time-domain response $x(t)$ is then

$$x(t) = \frac{\omega_0}{\sqrt{1-\zeta^2}} \cdot e^{-\zeta\omega_0 t} \cdot \sin\left(\omega_0\sqrt{1-\zeta^2}\right).$$

This is the response of a second-order sensor to a delta function input.

9.4.2 Linearly Increasing Input

Find a second-order sensor's response to the linearly increasing input $z(t) = kt$.

- The input's Laplace transform is

$$Z(s) = \mathcal{L}\{kt\} = \frac{k}{s^2},$$

and the sensor's output in the Laplace domain is then

$$X(s) = Z(s)\cdot H(s) = \frac{\omega_0^2}{s^2+2\zeta\omega_0 s+\omega_0^2}\cdot\frac{k}{s^2}$$

- As usual, we proceed by computing the partial fraction decomposition

$$X(s) = \frac{1}{s^2}\frac{k\omega_0^2}{s^2+2\zeta\omega_0 s+\omega_0^2} = k\omega_0^2\left(\frac{As+B}{s^2+2\zeta\omega_0 s+\omega_0^2}+\frac{Cs+D}{s^2}\right).$$

After forming a common denominator and combining like terms, we get

$$1 = (A+C)s^3 + (D+B+2\zeta\omega_0 C)s^2 + (C\omega_0^2 + 2\zeta\omega_0 D)s + D\omega_0^2,$$

and equating coefficients of powers of s produces the system of equations

$$\begin{aligned}
D\omega_0^2 = \omega_0^2 &\implies D = \frac{1}{\omega_0^2}\\
C\omega_0^2 + 2\zeta\omega_0 D = 0 &\implies C = -\frac{2\zeta}{\omega_0^3}\\
D+B+2\zeta\omega_0 C = 0 &\implies B = \frac{1}{\omega_0^2}\left(4\zeta^2-1\right)\\
A+C = 0 &\implies A = \frac{2\zeta}{\omega_0^3}.
\end{aligned}$$

The decomposition of $X(s)$, after cancelling ω_0^2 and rearranging slightly, is then

$$X(s) = k\left(\frac{1}{s^2} - \frac{2\zeta}{\omega_0}\cdot\frac{1}{s} + \underbrace{\frac{2\zeta}{\omega_0}\cdot\frac{s}{s^2+2\zeta\omega_0 s+\omega_0^2}}_{\equiv \mathrm{I}} + \underbrace{\frac{4\zeta^2-1}{s^2+2\zeta\omega_0 s+\omega_0^2}}_{\equiv \mathrm{II}}\right),$$

where we have defined the terms I and II for convenience; we will further rearrange individually. Before proceeding further, we also quote the identies

$$\mathcal{L}\left\{e^{at}\sin(\omega t)\right\} = \frac{\omega}{(s-a)^2+\omega^2} \tag{9.6}$$

$$\mathcal{L}\left\{e^{at}\cos(\omega t)\right\} = \frac{s-a}{(s-a)^2+\omega^2}. \tag{9.7}$$

- We first consider term I. Following the exact same procedure that led to Equation 9.4 in the previous problem, we get

$$\mathrm{I} = \frac{2\zeta}{\omega_0}\frac{s}{s^2 + 2\zeta\omega_0 s + \omega_0^2} = \frac{2\zeta}{\omega_0}\frac{s}{(s+\zeta\omega_0)^2 + \omega_0^2(1-\zeta^2)},$$

which, on comparison with the denominator in Equation 9.7, motivates the choice of terms $a = -\zeta\omega_0$ and $\omega^2 = \omega_0^2(1-\zeta^2)$. One more rearrangement then gives

$$\begin{aligned}\mathrm{I} &= \frac{2\zeta}{\omega_0}\frac{s+\zeta\omega_0}{(s+\zeta\omega_0)^2 + \omega_0^2(1-\zeta^2)} - \frac{2\zeta}{\omega_0}\frac{\zeta\omega_0}{(s+\zeta\omega_0)^2 + \omega_0^2(1-\zeta^2)} \\ &= \frac{2\zeta}{\omega_0}\frac{s-a}{(s-a)^2+\omega^2} - \frac{2\zeta^2}{(s-a)^2+\omega^2},\end{aligned}$$

where we have substituted in a and ω^2.

- Next, we consider the term II; we will also add on the term $-\frac{2\zeta^2}{(s-a)^2+\omega^2}$ term added on above. After writing the denomiator II in the familiar form $(s-a)^2+\omega^2$, we have

$$\begin{aligned}\mathrm{II} - \frac{2\zeta^2}{(s-a)^2+\omega^2} &= \frac{4\zeta^2-1}{(s-a)^2+\omega^2} - \frac{2\zeta^2}{(s-a)^2+\omega^2} = \frac{2\zeta^2-1}{(s-a)^2+\omega^2} \\ &= \frac{2\zeta^2-1}{\omega}\cdot\frac{\omega}{(s-a)^2+\omega^2}.\end{aligned}$$

The final form of the sensor's Laplace-domain respose $X(s)$ is then

$$X(s) = k\left(\frac{1}{s^2} - \frac{2\zeta}{\omega_0}\cdot\frac{1}{s} + \frac{2\zeta}{\omega_0}\cdot\frac{s-a}{(s-a)^2+\omega^2} + \frac{2\zeta^2-1}{\omega}\cdot\frac{\omega}{(s-a)^2+\omega^2}\right).$$

The corresponding time-domain response, using Equations 9.6 and 9.7, is

$$x(t) = k\left(t - \frac{2\zeta}{\omega_0} + \frac{2\zeta}{\omega_0}e^{-\zeta\omega_0 t}\cos(\omega t) + \frac{2\zeta^2-1}{\omega}e^{-\zeta\omega_0 t}\sin(\omega t)\right).$$

Finally, we note that for an optimally-damped sensor with $\zeta_{\text{opt}} = \frac{1}{\sqrt{2}}$, $x(t)$ reduces

$$x_{\text{opt}}(t) = k\left[t - \frac{\sqrt{2}}{\omega_0} + \frac{\sqrt{2}}{\omega_0}\exp\left(-\frac{\omega_0 t}{\sqrt{2}}\right)\cdot\cos\frac{\omega_0 t}{\sqrt{2}}\right]$$

The last term with $\sin\omega t$ is vanishes for $\zeta_{\text{opt}} = \frac{1}{\sqrt{2}}$. For $t\to\infty$, the optimally-damped sensor's output is

$$\lim_{t\to\infty} x(t) = k\left(t - \frac{2\zeta}{\omega_0}\right)$$

The sensor's output thus follows the input with a time offset $\frac{2\zeta}{\omega_0}$.

10 Tenth Exercise Set

10.1 TODO: A Sensor of Water Flow Speed

- Find equation of motion for the weight on the rod.

  Sum of torques:

  $$\sum_i M_i = I\ddot{\phi}$$

  Find all forces on the ball.

- Weight and buoyancy into effective gravitational acceleration g_{eff}.

  Consider only component of weight/bouyancy normal to the torque lever arm. This is

  $$mg_{\text{eff}} \sin\phi$$

- We find drag from the general equation

  $$F_v = 4\pi\eta r v$$

  Consider orbital velocity of body around origin. This is

  $$v_{\text{orb}} = \dot{\phi} l$$

  Note the fluid speed v_{f} moves "to the right" relative to a stationary frame. The projection of v_{f} normal to the lever arm (or onto v_{orb}) is $v_{\text{f}} \cos\phi$. Total speed normal to lever arm (which contribues via F_v to torque) is:

  $$F_v^{\perp} = 6\pi r\eta(v_{\text{f}} \cos\phi - v_{\text{orb}})$$

  This is component of F_v normal to the lever arm.

- Equation of motion is ($mg\sin\phi$ returns body to equilibrium position hence minus sign)

  $$I\ddot{\phi} = -mg\sin\phi l + F_v^{\perp} l = -mgl\sin\phi + 6\pi r\eta v_{\text{f}} \cos\phi - 6\pi r\eta l^2 \dot{\phi}$$

  odziv sensorja our output is ϕ while v is like input.

- We then write in sensor form

  $$I\ddot{\phi} + 6\pi r\eta l^2 \dot{\phi} + mgl\sin\phi = 6\pi r\eta v_{\text{f}} l \cos\phi$$

  Make small-angle approximation $\sin\phi \approx \phi$ and $\cos\phi \approx 1$ to get

  $$\ddot{\phi} + \frac{6\pi r\eta l^2}{I}\dot{\phi} + \frac{mgl}{I}\phi = \frac{6\pi r\eta l}{J} v_{\text{f}}.$$

  TODO: maybe change ϕ to theta or varphi.

  The above equation is the equation for a second-order sensor of the form

  $$\ddot{x} + 2\zeta\omega_0 \dot{x} + \omega_0^2 x = \omega_0^2 z + 2\zeta\omega_0 \dot{z}$$

  Note the factor $\frac{6\pi r\eta l^2}{I}$ also occurs the the RHS if we multiply above and below by l. We then get v_{f}/l corresponding to the sensor's output $\dot{z}$ from general second-order sensor equation. Note also there is no $\omega_0^2 z$ term in our water equation.

We then compare water equation term-by-term to the general second-order sensor equation to conclude

$$\omega_0^2 = \frac{mgl}{I} \qquad \text{and} \qquad 2\zeta\omega_0 = \frac{6\pi r\eta l^2}{I}$$

- Find parameters for optimal sensor (fastest reaction/following of input quantity). This requires $\zeta = \zeta_{\text{opt}} = \sqrt{2}$. In this case

$$\frac{1}{\sqrt{2}} \equiv \zeta = \frac{6\pi r\eta l^2}{I}\frac{1}{2\omega_0} = \frac{3\pi r\eta l^2}{I}\sqrt{\frac{I}{mgl}}$$

Alternatively square terms and get condition

$$mg = \frac{18\pi^2 r^2\eta^2 l^3}{I}$$

Mass, radius, length (and implicitly moment of inertia are parameters). We conclude

$$\frac{mJ}{r^2 l^3} = 18\pi^2\frac{\eta^2}{g}.$$

10.2 Theory: Sensor Response to Periodic Inputs

- We now consider sensors with periodic, sinusoidal inputs of the form

$$z(t) = z_0 e^{i\omega t} \tag{10.1}$$

where ω is the input signal's angular frequency and $z_0 \in \mathbb{R}$ is the input signal's (assumed to be real) amplitude. Often, when working with periodic inputs, we will use the term *filter* instead of *sensor*.

- Our goal is as follows: given a filter's transfer function $H(s)$ and an input signal $z(t)$, find the filter's output $x(t)$. In the Laplace domain, the relevant equation is

$$X(s) = H(s) \cdot Z(s).$$

- Claim: we can write an arbitrary transfer function in the form

$$H(s) = A\frac{P(s)}{D(s)}$$

where P and D are polynomials in s and $A \in \mathbb{R}$ is a real constant. Since P and D are polynomials, we may write $H(s)$ in the product form

$$H(s) = A\frac{\prod_a (s - s_a)}{\prod_b (s - s_b)}, \tag{10.2}$$

where s_a and s_b are the (in general complex) zeros of the polynomials $P(s)$ and $D(s)$ respectively. Thus s_a are the transfer function's *zeros* and s_b are the transfer function's *poles*. We then substitute Equation 10.2 into $X(s) = H(s)Z(s)$ and get

$$X(s) = H(s) \cdot Z(s) = \left(A\frac{\prod_a (s - s_a)}{\prod_b (s - s_b)}\right) \cdot \frac{z_0}{s - i\omega},$$

where we have assumed the periodic input $z(t) = z_0 e^{i\omega t}$ in Equation 10.1 to get the Laplace transform $Z(s)$. We proceed by writing $X(s)$ as a partial fraction decomposition, which reads

$$X(s) = Az_0 \left(\frac{\alpha_0}{s - i\omega} + \sum_b \frac{\alpha_b}{s - s_b} \right), \tag{10.3}$$

where the $\alpha_i \in \mathbb{C}$ are (in general complex) to-be-determined constants. Referring to Table 2, this expression then transforms back into the time domain as

$$x(t) = Az_0 \left(\alpha_0 e^{i\omega t} + \sum_b \alpha_b e^{s_b t} \right). \tag{10.4}$$

Recall the s_b are the transfer function $H(s)$'s poles, which are in general complex. We thus write the poles in the explicitly complex form $s_b = \sigma_b + i\omega_b$, where σ_b and ω_b are s_b's real and imaginary parts, respectively. Using $s_b = \sigma_b + i\omega_b$ this s_b, the terms in the above sum read

$$\alpha_b e^{s_b t} = \alpha_b e^{\sigma_b t} e^{i\omega_b t}.$$

This expression motivates the requirement $\sigma_b < 0$; otherwise $x(t)$ would diverge exponentially with time because of the $e^{\sigma_b t}$ term. Assuming $\sigma_b < 0$, we then write the sum terms as

$$\alpha_b e^{s_b t} = \alpha_b e^{-|\sigma_b| t} e^{i\omega_b t}.$$

- *Assumption*: we will consider sensor outputs for large times only, and ignore transient phenomena. In this case we have

$$e^{-|\sigma_b| t} \to 0 \qquad (\text{for } t \gg 1/|\sigma_b|),$$

in which the case the sum in Equation 10.4 vanishes, leaving only

$$x(t) = Az_0 \alpha_0 e^{i\omega t}.$$

We stress that this simplified expression for $x(t)$ neglects transient effects; however, transient phenomena often occur on such a short time scale that they can generally be neglected in non-specialized applications.

In any case, using $x(t) = Az_0\alpha_0 e^{i\omega t}$, we then write the sensor's output in the time domain as

$$x(t) = Az_0\alpha_0 e^{i\omega t} \equiv x_0 e^{i\omega t}, \tag{10.5}$$

where we have introduced the complex amplitude $x_0 \in \mathbb{C}$. Alternatively, in polar form, the complex amplitude reads

$$x_0 = |x_0| e^{i\delta},$$

where δ is the sensor's phase shift relative to the input. The sensor's output is then

$$x(t) = |x_0| e^{i\delta} e^{i\omega t}.$$

Next Step: Finding α_0

- We now return to the sensor's Laplace-domain response in Equation 10.3, which for review reads

$$X(s) = Az_0 \frac{\prod_a (s - s_a)}{\prod_b (s - s_b)} \cdot \frac{1}{s - i\omega} = Az_0 \left(\frac{\alpha_0}{s - i\omega} + \sum_b \frac{\alpha_b}{s - s_b} \right).$$

Next, following the standard partial fraction decomposition procedure, we multiply through by the denominator $\prod_b (s - s_b) \cdot (s - i\omega)$ and cancel the common amplitude term Az_0 to get

$$\prod_a (s - s_a) = \alpha_0 \prod_b (s - s_b) + (s - i\omega) \prod_b (s - s_b) \sum_b \frac{\alpha_b}{s - s_b}.$$

Our goal is to determine the constant term α_0, which corresponds to the periodic input $z(t) = z_0 e^{i\omega t}$. We can straightforwardly find α_0 by evaluating the above expression at $s = i\omega$. In this case most of the RHS vanishes, leaving only

$$\prod_a (i\omega - s_a) = \alpha_0 \prod_b (i\omega - s_b) \implies \alpha_0 = \frac{\prod_a (i\omega - s_a)}{\prod_b (i\omega - s_b)}$$

With α_0 known, the filter's time-domain response (recall Equation 10.5) is then

$$x(t) = Az_0 \alpha_0 e^{i\omega t} = Az_0 \frac{\prod_a (i\omega - s_a)}{\prod_b (i\omega - s_b)} e^{i\omega t} = A \frac{\prod_a (i\omega - s_a)}{\prod_b (i\omega - s_b)} z(t),$$

where we have recognized the original input signal $z(t) = z_0 e^{i\omega t}$. Comparing to the general expression $x(t) = H(s) \cdot z(t)$, we see the filter's transfer function is then

$$H(s) = A \frac{\prod_a (i\omega - s_a)}{\prod_b (i\omega - s_b)}.$$

However, noting that H is a function of only $i\omega$ and not the Laplace variable s, we write the filter's input-output relation (ignoring transient response) as

$$x(t) = H(i\omega) z(t).$$

Lesson: a sensor's transfer function in response to periodic input is just the general transfer function $H(s)$ with s replaced by $i\omega$.

Finally, we recall the polar notation $x(t) = |x_0| e^{i\delta} e^{i\omega t}$ to get

$$x(t) = |x_0| e^{i\delta} e^{i\omega t} = H(i\omega) \cdot z_0 e^{i\omega t} \implies H(i\omega) = \frac{|x_0|}{z_0} e^{i\delta}.$$

where z_0 is the sinusoidal input signal's (assumed to be) *real* amplitude.

Quantities We'll Be Interested In

- When analyzing filters with periodic inputs that may be written as $z(t) = z_0 e^{i\omega t}$, we will be interested in computing:

1. The ratio of the input and output signal's amplitudes, i.e.

$$A \equiv \frac{|x_0|}{z_0} = |H(i\omega)|. \tag{10.6}$$

   This quantity is called a filter's *gain* or *amplification*.

2. The phase shift δ of the output signal relative to the input signal, i.e.

$$\tan \delta = \frac{\mathrm{Im}\left\{H(i\omega)\right\}}{\mathrm{Re}\left\{H(i\omega)\right\}}. \tag{10.7}$$

10.3 Theory: Bode Plots

- Bode plots provide a graphical representation of a filter's gain $|H|$ (Equation 10.6) and phase shift δ (Equation 10.7) as a function of the frequency ω of a periodic input—we will consider only gain as a function of frequency.
- Bode plots are drawn in logarithmic scale to accomodate the large range of values for frequency and gain. Specifically, for Bode plots showing gain, we plot:
  - the quantity $20 \log |H|$ on the ordinate axis, and
  - the quantity $\log(\omega\tau) \equiv \log(\omega/\omega_c)$ on the abscissa.

  Note that both logarithms are base 10, note base e. **TODO:** what is it though? The parameter τ is the circuit's time constant (characteristic reponse time to periodic inputs?) and is used to define the *cut-off frequency* ω_c.
- Altough $20 \log |H|$ is technically a dimensionless quantity, it is assigned units of decibells (while $2 \log |H|$ would correspond to bells).
- **TODO:** provide a well-defined explanation of decibells.

10.4 Example: RC Low-Pass Filter

- **TODO:** RC element circuit.
- Define resistor R, capacitor C and input and output signals U_{in} and U_{out}, both measured relative to ground. Our first step is to find the circuit's transfer function, i.e. the relationship between input U_{in} and output U_{out}:

$$H(i\omega) = \frac{U_{\text{out}}}{U_{\text{in}}}$$

  For our analysis, we assume there is no current through the output node U_{out}, so any current through the resistor flows completely into the capacitor.
- We derive one equation from the voltage drop across the capacitor; this is

$$U_{\text{out}} - U_{\text{ground}} = U_{\text{out}} - 0 = Z_c \cdot I, \tag{10.8}$$

  where Z_c is the capacitor's complex impedance. We derive a second equation from the voltage drop across the entire circuit, which reads

$$U_{\text{in}} - U_{\text{ground}} = U_{\text{in}} - 0 = I(R + Z_c) \implies I = \frac{U_{\text{in}}}{R + Z_c}. \tag{10.9}$$

  In these equations we have explicitly written the ground voltage $U_{\text{ground}} = 0\,\text{V}$ for completeness, but we will drop this in the future.
- Next, we combine Equations 10.8 and 10.9 to find the circuit's transfer function:

$$U_{\text{out}} = Z_c I = Z_c \cdot \frac{U_{\text{in}}}{R + Z_c} \implies H(i\omega) = \frac{U_{\text{out}}}{U_{\text{in}}} = \frac{Z_c}{R + Z_c}.$$

  Finally, we substitute in the capacitor's impedance $Z_c = \frac{1}{i\omega C}$ to get

$$H(i\omega) = \frac{1}{1 + R/Z_c} = \frac{1}{1 + i\omega RC}.$$

- Finally, we note that, writing $i\omega \to s$, the filter's transfer function reads

$$H(s) = \frac{1}{1 + sRC},$$

which, after defining $\tau \equiv RC$, exactly matches the general first-order sensor form

$$H(s) = \frac{1}{1 + s\tau}.$$

Gain

- We now compute the circuit's gain from the transfer function's absolute value. From Equation 10.6, this is

$$|H|^2 = HH^* = \frac{1}{1 + i\omega\tau} \cdot \frac{1}{1 - i\omega\tau} = \frac{1}{1 + (\omega\tau)^2}$$

We then take the (positive) square root of $|H|^2$ to find the gain

$$A = |H| = \frac{1}{\sqrt{1 + (\omega\tau)^2}}.$$

- We now consider some limit cases. In the low-frequency limit $\omega\tau \ll 1$ the circuit's gain approaches

$$|H| \to 1 \qquad (\text{for } \omega\tau \ll 1).$$

In the high-frequency limit $\omega\tau \gg 1$ the circuit's gain approaches

$$|H| \approx \frac{1}{\omega\tau} \to 0 \qquad (\text{for } \omega\tau \gg 1).$$

The low- and high-frequency limits explain why the circuit is called a low-pass filter: low-frequency input signals are transmitted ("passed") unaffected with gain $A = 1$, while high-frequency inputs are attenuated or "blocked", since $A \to 0$ for large frequencies.

Bode Plot for a Low-Pass Filter

- Table 3 shows **TODO:** something clever

	$\lvert H\rvert$	$20 \log \lvert H\rvert$ [dB]
$\omega\tau \ll 1$	1	0
$\omega\tau \gg 1$	$1/(\omega\tau)$	$-20 \log(\omega\tau)$
$\omega\tau = 1$	$1/\sqrt{2}$	-3 dB

Table 3: Characteristic frequency regimes for a low-pass filter's Bode plot.

- Interpret the large-frequency ($\omega \gg \omega_c$) behavior of $20 \log |H|$, i.e. $-20 \log(\omega\tau)$ as a function of $\log(\omega\tau)$, as gain falling at 20 decibells per frequency decade.

  Recall bode plots are logarithmic, so plotting $-20 \log(\omega\tau)$ as a function of $\log(\omega\tau)$ in $\log(\omega\tau)$ space is equivalent to plotting the line $y = -20x$ on a linear xy scale.

- Terminology: negative gain is also called attenuation.

11 Eleventh Exercise Set

11.1 Ideal Integration Circuit

- An ideal integration circuit takes an input signal U_{in} and outputs a signal U_{out} obeying

$$U_{\text{out}}(t) = k \int_0^t U_{\text{in}}(t')\,\mathrm{d}t'. \tag{11.1}$$

To find the circuit's transfer function, we first take the time derivative to get

$$\dot{U}_{\text{out}}(t) = kU_{\text{in}}(t).$$

We then take the above equation's Laplace transform and, assuming $U_{\text{in}}(0) = 0$, get

$$sU_{\text{out}}(s) = kU_{\text{in}}(s) \implies H_{\text{int}}(s) = \frac{U_{\text{out}}}{U_{\text{in}}} = \frac{k}{s}.$$

For periodic input signals, where we can replace s with $i\omega$, an ideal integrator's transfer function reads

$$H_{\text{int}}(i\omega) = \frac{k}{i\omega}. \tag{11.2}$$

- Next, we recall a RC low-pass filter has a transfer function **TODO:** reference

$$H_{\text{lpf}}(i\omega) = \frac{1}{1 + i\omega\tau} \overset{\omega\tau \gg 1}{\longrightarrow} H_{\text{lpf}}(i\omega) \approx \frac{1}{i\omega\tau}.$$

In other words, for high frequencies $\omega\tau \gg 1$, a low-pass filter behaves like an ideal integrator (Equation 11.1) with a constant $k = \frac{1}{\tau}$. However, an RC low-pass filter is not a particularly useful integrator. Here is the problem: for $\omega\tau \gg 1$, i.e. the regime in which a LPF behaves as an integrator, the filter also exponentially attenuates the input signal, so the integration circuit's output would be essentially zero, and obscured by thermal noise.

11.2 Example: CR High-Pass Filter

- See notes for diagram. Define current I through C to R to ground. Voltage drop across resistor is $U_{\text{out}} = RI$. Voltage drop across circuit is $U_{\text{in}} = I(Z_{\text{C}} + R)$. We equate currents in both expressions to get

$$\frac{U_{\text{out}}}{R} = I = \frac{U_{\text{in}}}{Z_{\text{C}} + R}$$

We then solve for the transfer function $H(i\omega)$ and get

$$\begin{aligned} H(i\omega) \equiv \frac{U_{\text{out}}}{U_{\text{in}}} &= \frac{R}{Z_{\text{C}} + R} = \frac{R}{1(i\omega C) + R} = \frac{i\omega RC}{1 + i\omega RC} \\ &\equiv \frac{i\omega\tau}{1 + i\omega\tau}, \end{aligned}$$

where we have defined the time constant $\tau \equiv RC$.

- We find the circuit's gain $|H|$ from is

$$|H|^2 = \frac{i\omega\tau}{1 + i\omega\tau} \cdot \frac{(-i)\omega\tau}{1 + (-i)\omega\tau} \implies |H| = \frac{\omega\tau}{\sqrt{1 + (\omega\tau)^2}}.$$

	$\lvert H\rvert$	$20\log\lvert H\rvert$ [dB]
$\omega\tau \ll 1$	$\omega\tau$	$20\log(\omega\tau)$
$\omega\tau \gg 1$	1	0
$\omega\tau = 1$	$1/\sqrt{2}$	$-3\,\mathrm{dB}$

Table 4: Characteristic frequency regimes for a high-pass filter's Bode plot.

- The high-pass filter's amplification behavior in three characteristic frequency regimes is shown in Table 4.

  TODO: plot Bode plot: gain $|H|$ increases linearly at 20 decibells per frequency decade in the regime $\omega\tau \ll 1$, reaches $|H| = -3\,\mathrm{dB}$ at $\omega\tau = 1$ and remains constant for $\omega\tau \gg 1$.

  The circuits amplification behavior shows why it is called a high-pass filter—low frequencys for $\omega \ll \omega_c$ are strongly attenuated, while high frequencies with $\omega \gg \omega_c$ are passed with gain one.

- Finally, we note the interpretation of a high-pass filter as a first-order sensor with the transfer function

$$H_{\mathrm{hpf}}(s) = \frac{\tau s}{1+\tau s}, \quad \text{where } \tau = RC.$$

11.3 Ideal Differentiator

- An ideal differentiation circuit takes an input signal U_{in} and outputs a signal U_{out} obeying

$$U_{\mathrm{out}}(t) = k\dot{U}_{\mathrm{in}}(t). \tag{11.3}$$

  To find the circuit's transfer function, we first transform to the Laplace domain and, assuming $U_{\mathrm{in}}(0) = 0$, get

$$U_{\mathrm{out}}(s) = ksU_{\mathrm{in}}(s) \implies H_{\mathrm{dif}}(s) = \frac{U_{\mathrm{out}}}{U_{\mathrm{in}}} = ks.$$

  For periodic inputs, writing $s \to i\omega$, the ideal differentiator's transfer function reads

$$H_{\mathrm{dif}}(i\omega) = k \cdot i\omega. \tag{11.4}$$

- Next, recall **TODO:** reference a high-pass filter has a transfer function

$$H_{\mathrm{hpf}}(i\omega) = \frac{i\omega\tau}{1+i\omega\tau} \overset{\omega\tau\ll 1}{\longrightarrow} H_{\mathrm{hpf}}(i\omega) \approx i\omega\tau.$$

  In other words, a for low-frequency inputs $\omega \ll \omega_c$, a high-pass filter behaves as an ideal differentiator (Equation 11.3) with constant $k = \tau$. However, like the LPF when used as an integrator, a HPF is not a useful differentiator, since a HPF strongly attenuates its input signal in the differentiation regime $\omega \ll \omega_c$.

11.4 Theory: Active Circuits and the Operational Amplifier

- Active circuit elements are capable of supplying energy to a circuit, while passive elements cannot supply energy. A circuit containing an active circuit element is called an active circuit; circuits with only passive elements are passive circuits.

- Resistors, capacitors and inductors are examples of passive elements, while batteries/power supplies are active elements. We will now introduce another important active element called an *operational amplifier*, or simply op-amp.

Review of Op-Amp

- An op-amp has two inputs: the non-inverting input U_+ and the inverting input U_-.

  An op-amp's transfer function reads

  $$U_{\text{out}} = A(s)(U_+ - U_-), \tag{11.5}$$

  where $A(s)$ is the amplifier's (in general frequency-dependent) gain.

- Table 5 summarizes the op-amps important properties

Ideal Op-Amp	Real Op-Amp
$A(s) \equiv A_0 = \text{constant}$	$A(s) = A_0 \cdot H_{\text{lpf}}(s)$
$A_0 \to \infty$	$A_0 \sim 10^4$ to 10^6
$Z_+, Z_- \to \infty$	$Z_+, Z_- \sim 10^{10}$ to $10^{12}\,\Omega$
$I_+, I_- \to 0$	$I_+, I_- \lesssim 10^{-12}\,\text{A}$
$Z_{\text{out}} \to 0$	$Z_{\text{out}} \sim 10^{-3}\,\Omega$

Table 5: Comparison of an ideal and real-life operational amplifier's important properties.

Negative Feedback Loop

- **TODO:** see notes for schematic.

  Input signal U_{in} and op-amp. Input signal led into op-amps non-inverting input, and op-amp's output U_{out}.

  In addition to serving as the circuit's output, is also led back into the op-amp's inverting input in a *negative feedback loop* through an arbitrary circuit with transfer function $F(s)$.

- The input to the op-amp's inverting input is thus $F(s) \cdot U_{\text{out}}$, while the op-amp's output, referring to Equation 11.5, is

  $$U_{\text{out}} = A(s) \cdot \big(U_{\text{in}} - F(s) \cdot U_{\text{out}}\big).$$

  We assume the op-amp is ideal, so that $A \equiv A_0$, and rearrange to get

  $$U_{\text{out}} \cdot (1 + A_0 F) = A_0 U_{\text{in}} \implies H \equiv \frac{U_{\text{out}}}{U_{\text{in}}} = \frac{A_0}{1 + A_0 F} = \frac{1}{(1/A_0) + F}.$$

  In the ideal op-amp limit $A_0 \to \infty$, the circuit's transfer function simplifies to

  $$H(s) = \lim_{A_0 \to \infty} \frac{1}{(1/A_0) + F} = \frac{1}{F(s)}.$$

  Importantly, the circuit's transfer function is independent of the op-amp's gain A. This is desirable, since, in practice, A is an unpredictable parameter.

11.5 Examples of Op-Amp Circuits

11.5.1 Differentiator

- **TODO:** see notes for circuit.
- Input signal U_{in} flows through capacitor C and branches. One branch into op-amps U_- and one branch through op-amp negative feedback loop through a resistor R before connecting to U_{out}. Non-inverting input U_+ is grounded.
- Assume current through capacitor is I and apply ideal op-amp properties. First: potential at U_- will match ground potential at U_+; second: current into U_- is zero. Combining these gives
$$U_{\text{in}} = IZ_{\text{C}}.$$
No current flows into U_-, so any current through C must flow through R, giving
$$-U_{\text{out}} = IR.$$
We then equate currents to get
$$U_{\text{out}} = -\frac{R}{Z_{\text{C}}}U_{\text{in}} \implies H = \frac{U_{\text{out}}}{U_{\text{in}}} = -i\omega RC \equiv -i\omega\tau.$$
This matches the expression for an ideal differentiator (for periodic inputs) in Equation 11.4.

11.5.2 Integrator

- **TODO:** see notes for circuit.
- Input signal U_{in} flows through resistor R and branches. One branch into op-amp's U_- and one branch through op-amp negative feedback loop through a capacitor C before connecting to U_{out}. Non-inverting input U_+ is grounded.
- Assume current through resistor is I and apply ideal op-amp properties. First: potential at U_- will match ground potential at U_+; second: current into U_- is zero. Combining these gives
$$U_{\text{in}} = IR.$$
No current flows into U_-, so any current through R must flow through C, giving
$$-U_{\text{out}} = IZ_{\text{C}}.$$
We then equate currents to get
$$U_{\text{out}} = -\frac{Z_{\text{C}}}{R}U_{\text{in}} \implies H = \frac{U_{\text{out}}}{U_{\text{in}}} = -\frac{1}{i\omega RC} \equiv -\frac{1}{i\omega\tau}.$$
This, assuming $k = -1/(RC)$, matches the expression for an ideal integrator (for periodic inputs) in Equation 11.2

12 Twelfth Exercise Set

12.1 Circuits Using Op-Amps

12.1.1 Inverting Amplifier

- **TODO:** see notes for circuit.
- Input signal U_{in} flows through resistor R_{in} and branches. One branch into op-amp's U_- and one branch through op-amp negative feedback loop through a resistor R_{f} before connecting to U_{out}. Non-inverting input U_+ is grounded.
- Assume current through R_{in} is I and apply ideal op-amp properties. First: potential at U_- will match ground potential at U_+; second: current into U_- is zero. Combining these gives

  $$U_{\text{in}} = IR_{\text{in}}.$$

  No current flows into U_-, so any current through R_{in} must flow through R_{f}, giving

  $$-U_{\text{out}} = IR_{\text{f}}.$$

  We then equate currents to get

  $$U_{\text{out}} = -\frac{R_{\text{f}}}{R_{\text{in}}}U_{\text{in}} \implies H = \frac{U_{\text{out}}}{U_{\text{in}}} = -\frac{R_{\text{f}}}{R_{\text{in}}}.$$

- Interpretation: an inverting amplifier's gain is $G = -R_{\text{f}}/R_{\text{in}}$—because the amplifier reverses the sign of the input signal, it is called an inverting amplifier.

12.1.2 Non-Inverting Amplifier

- **TODO:** see notes for circuit.
- Input signal U_{in} flows directly into non-inverting input U_+.

  Grounded resistor R_{g}, branch, one branch into inverting input U_- and one through negative feed back loop via resistor R_{f} connecting to op-amp output U_{out}.
- Assume current flows from op-amp output backward through R_{f} and apply ideal op-amp properties. First: potential at U_- will match input potential U_{in} at the non-inverting input U_+; second: current into U_- is zero, so all current through R_{f} must flow via R_{g} into ground. Combining these gives

  $$U_{\text{out}} - U_{\text{in}} = IR_{\text{f}} \qquad \text{and} \qquad U_{\text{in}} = IR_{\text{g}}.$$

  We then equate currents to get

  $$U_{\text{out}} - U_{\text{in}} = \frac{R_{\text{f}}}{R_{\text{g}}}U_{\text{in}} \implies H = \frac{U_{\text{out}}}{U_{\text{in}}} = 1 + \frac{R_{\text{f}}}{R_{\text{g}}}.$$

- Interpretation: a non-inverting amplifier's gain is $G = 1 + R_{\text{f}}/R_{\text{in}}$—because the amplifier preserves the input signal's sign, it is called a non-inverting amplifier.

12.1.3 Voltage Follower

- **TODO:** see notes for circuit.
- A voltage follower is just an op-amp with an "empty" negative feedback loop. Neglecting transient effects, the input-out relationship is simply

$$U_{\text{out}} = U_{\text{in}} \implies H = 1.$$

  Although it might appear pointless at first glance, a voltage follower is useful because of its high input impedance (very little current flows into the op-amp) and low output impedance.

  A voltage follower can be used to isolate two sub-circuits such that no current flows between them. Example follows

Example Use of Voltage Follower

- **TODO:** see notes for circuit.
- Essentially two coupled RC low-pass filters with values R_1, C_1 and R_2, C_2; both capacitors are grounded. Input signal U_{in} flows through R_1, output from R_1 is input to R_2, and output of R_2 is U_{out}.

  Note that the circuit's output is not simply $H_{\text{tot}} = H^{(1)}_{\text{lpf}} \cdot H^{(2)}_{\text{lpf}}$. This won't work because our derivation of the LPF transfer function assumed no current flows through its output; this clearly does not hold for the first RC filter.
- Solution: place a voltage follower between the two LPF such that R_1's output serves as the input into the voltage follower's non-inverting input U_+ and the voltage follower's output serves as the input into R_2. The voltage follower's high input impedance means practically no current flows out of the first RC filter, which justifies $H_{\text{tot}} = H^{(1)}_{\text{lpf}} \cdot H^{(2)}_{\text{lpf}}$.

12.2 TODO: A Differential Amplifier

- Note that this is not the most common op-amp differential amplifier.
- First note U_1 between R_1 and R_2. Similarly note U_2 between R_2' and R_1'. Let x denote voltage between R_2 and R_2'. Assume no current into opamp minus inputs, so current flows completely from R_1 to R_2 and from R_2' and R_1' (and note into non-inverting inputs).
- Begin by analyzing currents. Start with primed (bottom, near ground) subcircuit. Assume current I_2 from output of opamp 2. Similarly, assume current I_1 from output of opamp 1. The top current must flow fully through R_1 (since not input U_{out}).
- In principle, current I_1 could also flow up through R_2!

  However, we now the same current must flow through R_2' and R_1'. First condition is $I_{R_2'} = I_{R_1'} \equiv I'$. We thus write

$$I' = \frac{x - U_2}{R_2'} = \frac{U_2}{R_1'}.$$

We solve for x to get

$$x = U_2 \left(1 + \frac{R_2'}{R_1'}\right).$$

Note this is the same behavior of a non-inverting amplifier. (this follows from the conclusin that $I_{R_2'} = I_{R_1'}$).

- Next we note that the same current must flow through R_1 and R_2 (even through we don't know what it is). We then have $I_{R_2} = I_{R_1} \equiv I$. We thus have

$$I = \frac{U_{\text{out}} - U_1}{R_1} = \frac{U_1 - x}{R_2}.$$

We can then solve for U_{out} to get

$$U_{\text{out}} = U_1 \left(1 + \frac{R_1}{R_2}\right) - x \cdot \frac{R_1}{R_2}.$$

Finally, we can substitute in x from Equation **TODO:** reference to get

$$U_{\text{out}} = U_1 \left(1 + \frac{R_1}{R_2}\right) - \frac{R_2}{R_2} U_2 \left(1 + \frac{R_2'}{R_1'}\right)$$

- Next step is to find the ratio of voltages for which

$$U_{\text{out}} = k(U_1 - U_2),$$

i.e. for the circuit to act as a subtraction circuit. The condition comes out to the coefficients of U_1 and U_2 being equal; this reads

$$1 + \frac{R_1}{R_2} = \frac{R_1}{R_2}\left(1 + \frac{R_2'}{R_1'}\right) = \frac{R_1}{R_2} + \frac{R_1}{R_2} \cdot \frac{R_2'}{R_1'}$$

The results is

$$\frac{R_1}{R_2} \cdot \frac{R_2'}{R_1'} = 1 \implies \frac{R_1}{R_2} = \frac{R_1'}{R_2'}.$$

Under this condition the circuit's output reads

$$U_{\text{out}} = \left(1 + \frac{R_1}{R_2}\right) \cdot (U_1 - U_2).$$

Note that the coefficient serves as an amplification coefficient. The trasnfer function would reads

$$H = \frac{U_{\text{out}}}{U_1 - U_2}$$

Careful noting that the circuit's input voltage is $U_1 - U_2$.

12.3 TODO: An Amplifying Feedback Loop

- General feedback loop with an arbitrary transfer function $H(s)$. Then a triangle written with K that serves as an amplifier with gain K. Goal is to find the value of K such that, for a linearly-increasing input $U = \alpha t$, the circuit follows the inputs derivative three times as fast as it follows the circuit itself.

- We are given
$$H(s) = \frac{\tau s}{1+\tau s}.$$

A First Look Without the Feedback Loop

- In this case we consider just U_{in} through $H(s)$, which outputs U_{out}. We begin with the Laplace transformation
$$U_{\text{out}}(s) = H(s)U_{\text{in}}(s) = \frac{\tau s}{1+\tau s} \cdot \frac{\alpha}{s^2}$$
Solve with partial fraction decomposition and get
$$U_{\text{out}}(s) = \alpha\tau\left(\frac{A}{s} + \frac{B}{1+\tau s}\right) \implies 1 + s(\tau A + B) + A$$
From this we conclude
$$A = 1 \qquad \text{and} \qquad B = -\tau A = -\tau,$$
in terms of which the circuit's Laplace-domain output equals
$$U_{\text{out}}(s) = \alpha\tau\left(\frac{1}{s} - \frac{1}{s+(1/\tau)}\right)$$

- From this circuit's time-domain output is
$$U_{\text{out}}(t) = \alpha\tau\left(1 - e^{-t/\tau}\right)$$
From this we see that in the limit $t \gg \tau$ we have $U_{\text{out}}(t) \to \alpha\tau \neq U_{\text{in}}(t) = \alpha t$.
In this case, up to a multiplicative constant, the circuit acts like a differentiation circuit since $\alpha\tau \propto \frac{\mathrm{d}U_{\text{in}}}{\mathrm{d}t} = \alpha$.
However, we have a problem: as we speed up the circuit's response by decreasing τ, the circuit's output falls in amplitude, and would eventually be comparable to noise.

With the Feedback Loop

- In this case consider output voltage U_{out}, which is also input into the gain circuit in the feedback loop, whose output is KU_{out}.
The input into the $H(s)$ circuit is thus $U_{\text{in}} - KU_{\text{out}}$, so the output is
$$U_{\text{out}} = H(s)(U_{\text{in}} - KU_{\text{out}}).$$

- To review, the goal is to find the entire circuit's transfer function
$$H_{\text{tot}}(s) = \frac{U_{\text{out}}}{U_{\text{in}}}.$$
We begin with
$$U_{\text{out}} = H(s)(U_{\text{in}} - KU_{\text{out}}).$$

and rearrange to get

$$U_{\text{out}}(1+KH) = HU_{\text{in}} \implies H_{\text{tot}} = \frac{U_{\text{out}}}{U_{\text{in}}} = \frac{H}{1+KH}.$$

From the given transfer function $H = \frac{\tau s}{1+\tau s}$ we see that, after some straightforward, algebra the total transfer function

$$H_{\text{tot}}(s) = \frac{\tau s}{1+\tau s+K\tau s} = \frac{\tau s}{1+(1+K)\tau s}.$$

- We now aim to express H_{tot} in a way it will be clear that the circuit with the feedback loop still acts as a differentiator. We have

$$H_{\text{tot}} = \frac{1}{1+K}\frac{(1+K)\tau s}{1+(1+K)\tau s} \equiv \frac{1}{1+K}\frac{\widetilde{\tau} s}{1+\widetilde{\tau} s}.$$

Our condition for three-times faster following of derivative in feedback loop mode is

$$\widetilde{\tau} = \frac{\tau}{3} \implies (1+K)\tau = \frac{\tau}{3} \implies K = -\frac{2}{3}.$$

- We now see if adding the feedback loop solves the problem of output amplitude falling as the circuit's response increases. We then have (adapting previous non-feedback results), the output

$$U_{\text{out}}(t) = \frac{1}{1+K}\alpha\widetilde{\tau}(1-e^{-t/\widetilde{\tau}}).$$

Noting that $\widetilde{\tau} = (1+K)\tau$, we then have

$$U_{\text{out}}(t) = \alpha\tau(1-e^{-t/\tau}).$$

Lesson: the amplitude of the circuit's output is not any smaller than without the feedback loop, the circuit's response is three times faster.

12.4 Band-Pass Filter

- **TODO:** see notes for circuit.
- Consider input signal U_{in} through resistor R. Branch. Right branch to U_{out} and bottom branch to an additional CL branch with capacitor C and inductor L, which then region and pass into ground.
- Define an equivalent circuit in which the CL tree is replaced by an element with impedance Z_{eq}.
- Assume current out of U_{out} is zero, so all current through R flows entirely through Z_{eq} and into ground. We then write the voltage drop equations

$$U_{\text{in}} - U_{\text{out}} = IR \qquad \text{and} \qquad U_{\text{out}} = IZ_{\text{eq}}$$

and equate currents to get

$$\frac{U_{\text{in}}-U_{\text{out}}}{R} = \frac{U_{\text{out}}}{Z_{\text{eq}}} \implies U_{\text{in}} = U_{\text{out}}\left(1+\frac{R}{Z_{\text{eq}}}\right).$$

We then solve for the circuit's transfer function, which comes out to

$$H = \frac{U_{\text{out}}}{U_{\text{in}}} = \frac{Z_{\text{eq}}}{Z_{\text{eq}}+R}.$$

- We will now find an explicit expression for the equivalent impedance Z_{eq}, which we compute from

$$Z_{\text{eq}}^{-1} = Z_{\text{C}}^{-1} + Z_{\text{L}}^{-1} = i\omega C + \frac{1}{i\omega L} = \frac{1-\omega^2 LC}{i\omega L} \implies Z_{\text{eq}} = \frac{i\omega L}{1-\omega^2 LC}.$$

  In terms of the just-computed equivalent impedance Z_{eq}, the circuit's transfer function, after forming common denominators, is

$$H = \frac{i\omega L}{R - \omega^2 LCR + i\omega L}.$$

- Next, to match convention, we will write the transfer function in terms of the complex frequency $s = i\omega$, i.e.

$$H(s) = \frac{sL}{R + s^2 LCR + sL}.$$

  We then divide numerator and denominator by LRC to get

$$H(s) = \frac{s/(RC)}{s^2 + s/(RC) + 1/(LC)}.$$

  We then introduce the new variables $\omega_0^2 \equiv 1/(LC)$ (this is the undamped circuit's resonance frequency) and $\omega_{\text{c}} = 1/(RC)$, in terms of which the circuit's transfer function reads

$$H(s) = \frac{s\omega_{\text{c}}}{s^2 + s\omega_{\text{c}} + \omega_0^2}.$$

Interpretation

- We analyze the circuit's output by finding the transfer function's zero and poles. The only zero is the trivial zero $s = 0$, while the transfer function's poles are the solution to the quadratic equation

$$s^2 + s\omega_{\text{c}} + \omega_0^2 = 0$$

  Using the quadratic formula, the poles are

$$s_\pm = -\frac{\omega_{\text{c}}}{2} \pm \frac{1}{2}\sqrt{\omega_{\text{c}}^2 - 4\omega_0^2} = -\frac{\omega_{\text{c}}}{2} \pm \frac{i}{2}\sqrt{4\omega_0^2 - \omega_{\text{c}}^2}.$$

  We will consider the limit case $\omega_0 \gg \omega_{\text{c}}$, in which case the filter's range of passed frequencies is "narrow" (relative to?). Assuming $\omega_0 \gg \omega_{\text{c}}$ the two poles read

$$s_\pm = -\frac{\omega_{\text{c}}}{2} \pm \frac{i}{2}\sqrt{4\omega_0^2 - \omega_{\text{c}}^2} \approx -\frac{\omega_{\text{c}}}{2} \pm i\omega_0.$$

Theory: Pole-Zero Plot

- We draw s in the complex plane with $\operatorname{Re} s$ on the x axis and $\operatorname{Im} s$ on the abscissa. Zeros are drawn with a cross $\times$ symbol and poles with a circle $\circ$ symbol. Without proof, the pole closest to ω_0 turns out to have the strongest effect on the transfer function's behavior.

Next

- Goal is to find width of band pass filter for as a function of ω_0 and ω_c in a neighborhood of ω_0. We quantify the band pass filter's width as the frequencies at which the BPF's $|H(\omega)|$ characteristic falls by $-3\,\mathrm{dB}$ relative to the maximum resonance value $|H|(\omega_0)$.

- Write factored transfer function as

$$H(s) = \frac{s\omega_c}{(s - s_+)(s - s_-)} \to \frac{i\omega\omega_c}{(i\omega - s_+)(i\omega - s_-)}$$

- If we draw $\pm\omega_0$ on the $\mathrm{Im}\, s$ axis, we find that $s_+ = -(\omega_c/2) + i\omega_0$ is much closer to the resonance frequency ω_0 than s_-, especially in the limit $\omega_0 \gg \omega_c$.

  We thus consider only the effect of s_+ on the transfer function, and introduce the new transfer function

$$\widetilde{H} = \frac{1}{s - s_+} = \frac{1}{i(\omega - \omega_0) + (\omega_c/2)}.$$

  This transfer function is a good approximation for $H(s)$ near the resonance frequency ω_0 in the limit $\omega_0 \gg \omega_c$.

13 Thirteenth Exercise Set

13.1 Band-Pass Filter (continued)

- Begin with approximation to transfer function for $\omega_0 \gg \omega_c$.

$$\widetilde{H} = \frac{1}{s - s_+} = \frac{1}{i\omega - i\omega_0 + \omega_c/2} \implies |\widetilde{H}| = \frac{1}{\sqrt{(\omega - \omega_0)^2 + (\omega_c/2)^2}}$$

Our goal is to find width of resonance peak $|\widetilde{H}|(\omega)$ at an attenuation of $-3\,\mathrm{dB}$. In other words, we aim to find the frequencies solving the equation

$$20 \log \left(\frac{|\widetilde{H}|}{|\widetilde{H}|_{\max}} \right) \mathrm{dB} = -3\,\mathrm{dB}.$$

To avoid problems with square root in definition of $|\widetilde{H}|$ we take use logarithm identity

$$10 \log \left(\frac{|\widetilde{H}|^2}{|\widetilde{H}|^2_{\max}} \right) \mathrm{dB} = -3\,\mathrm{dB}. \tag{13.1}$$

We then make the auxiliary calculation

$$|\widetilde{H}|^2_{\max} = \left(|\widetilde{H}|(\omega_0) \right)^2 = \frac{1}{(\omega_c/2)^2} = \frac{4}{\omega_c^2},$$

which we substitute into Equation 13.1 and divide through by 10 to get

$$\log \left[\frac{\omega_c^2}{4} \cdot \frac{1}{(\omega - \omega_0)^2 + (\omega_c/2)^2} \right] = -\frac{3}{10} = -0.3.$$

We then exponentiate both sides with base 10 to eliminate the logarithm, leaving

$$\frac{\omega_c^2}{4} \cdot \frac{1}{(\omega - \omega_0)^2 + (\omega_c/2)^2} = -\frac{3}{10} = 10^{-0.3} = 0.501 \approx \frac{1}{2}.$$

We then solve the resulting polynomial equation for ω, which results it

$$(\omega - \omega_0)^2 = \frac{\omega_c^2}{4} \implies \omega_\pm = \omega_0 \pm \frac{\omega_c}{2}.$$

The band pass filter's frequency band, i.e. the frequency range in which the filter's gain is greater than $-3\,\mathrm{dB}$, is thus

$$\Delta\omega_{-3\,\mathrm{dB}} = \omega_+ - \omega_- = \omega_c.$$

Quality Factor

- In passing, we note that a resonance filter's *quality factor* Q is defined as

$$Q \equiv \frac{\omega_0}{\Delta\omega_{-3\,\mathrm{dB}}},$$

which in our case comes out to $Q = \omega_0/\omega_c$. The narrower a filter's resonance peak, the larger the filter's quality factor. In other words, a high-quality filter only passes a very narrow frequency range.

- Note that for the above band pass filter, for which $\omega_0^2 = 1/(LC)$ and $\omega_c = 1/(RC)$ we have

$$Q = \frac{\omega_0}{\omega_c} = \frac{RC}{\sqrt{LC}}.$$

13.2 Band-Stop Filter

- **TODO:** see notes for circuit.
- Input signal U_{in} passes through resistor R and branches. First branch passes through a capacitor C and inductor L, wired in series, and then into ground; the second branch is the filter's output U_{out}.
- We model the CL series branch as a single equivalent circuit with impedance

$$Z_{\text{eq}} = Z_{\text{C}} + Z_{\text{L}} = \frac{1}{i\omega C} + i\omega L = \frac{1}{sC} + sL = \frac{1 + s^2 LC}{sC},$$

where we have replaced $i\omega$ with the Laplace-domain frequency s. We then write the voltage drop equations

$$U_{\text{in}} - U_{\text{out}} = IR \qquad \text{and} \qquad U_{\text{out}} = IZ_{\text{eq}}$$

and then equate currents to get

$$U_{\text{in}} - U_{\text{out}} = \frac{R}{Z_{\text{eq}}} U_{\text{out}} \implies H = \frac{U_{\text{out}}}{U_{\text{in}}} = \frac{Z_{\text{eq}}}{R + Z_{\text{eq}}}.$$

After substituting in the equivalent impedance, the filter's transfer function is

$$H(s) = \frac{1 + s^2 LC}{sRC + 1 + s^2 LC}.$$

We then divide numerator and denominator by LC to get the standard expression

$$H(s) = \frac{s^2 + (1/LC)}{s^2 + (R/L) \cdot s + 1/(LC)} \equiv \frac{s^2 + \omega_0^2}{s^2 + s\omega_{\text{c}} + \omega_0^2},$$

where we have defined the parameters $\omega_{\text{c}} = R/L$ and $\omega_0^2 = 1/(LC)$.

- Following an identical procedure as for a band-pass filter in the previous exercise, we find that the band-stop filter's frequency width at $-3\,\text{dB}$ attenuation is

$$\Delta\omega_{-3\,\text{dB}} = \omega_{\text{c}} = \frac{R}{L}.$$

The circuit's quality is

$$Q \equiv \frac{\omega_0}{\Delta\omega_{-3\,\text{dB}}} = \frac{\omega_0}{\omega_{\text{c}}} = \frac{L}{R}\frac{1}{\sqrt{LC}}.$$

13.3 Theory: Statistics

- Consider a set of N data points $\{z_i\}_{i=1}^{N}$, for example measurements of a physical quantity. We will call the set of points $\{z_i\}$ a *sample*.

  Statistics assumes the sample points are drawn from a *population* described by a probability distribution $\frac{\text{d}p}{\text{d}z}$; in the scope of this course, we will assume the sample values z_i are from a normally-distributed population, and thus obey $z_i \sim \mathcal{N}(\mu, \sigma^2)$.

  Importantly, the normal distribution's mean μ or variance σ^2 are in general unknown. One goal of statistics is to estimate the parameters of the distribution describing a sample of data points—in our case these parameters would be μ and σ^2.

- Functions of the sample values z_i are called *sample statistics*, and can be used as *estimators* of the population parameters μ and σ^2.

Estimator of Population Mean

- The estimator function for the population mean μ is just the average of the sample values $\{z_i\}$ and reads

$$\overline{z} = \frac{1}{N}\sum_{i=1}^{N} z_i.$$

- Formally, the factor $1/N$ comes from the requirement that the sample mean estimator $\overline{z}$ be *unbiased*—this means that the expectation value of the estimator must equal the quantity it estimates (i.e. $\langle \overline{z} \rangle = \mu$).

  To derive the coefficient $1/N$, we first write the mean estimator in the general form

$$u = k\sum_{i=1}^{N} z_i,$$

  where k is to be determined such that u is an unbiased estimator of population mean μ. The unbiased condition reads

$$\langle u \rangle = k\left\langle \sum_{i}^{N} z_i \right\rangle = k\sum_{i=1}^{N} \langle z_i \rangle = kN\mu \implies k = \frac{1}{N},$$

  where we have used the identity $\langle z_i \rangle = \mu$, since the sample values are assumed to be normally distributed about the mean value μ. Requiring $K = 1/N$, the estimator of the population mean is then

$$u = \frac{1}{N}\sum_{i=1}^{N} z_i \equiv \overline{z}.$$

Estimator of Variance

- For a sample $\{z_i\}$ drawn for a population distributed as $\mathcal{N}(\mu, \sigma^2)$, the estimator of the population variance σ^2 is

$$s^2 = \frac{1}{N-1}\sum_{i=1}^{N} (z_i - \overline{z})^2, \tag{13.2}$$

  where $\overline{z}$ is the sample mean. Without derivation, the coefficient $1/(N-1)$ comes from the requirement that $\langle s^2 \rangle = \sigma^2$, i.e. that s^2 is an unbiased estimator of σ^2.

Comment: Estimating Variance if Mean is Known

- So far, for generality, we have assumed both the population mean μ and population variance σ^2 are unknown. Now, hypothetically assume we know the mean μ of the population from which a sample is drawn. In this case we define a new estimator for σ^2 according to

$$w^2 = k\sum_{i=1}^{N} (z_i - \mu)^2.$$

We then apply the unbiased requirement $\langle w^2 \rangle = \sigma^2$ to get

$$\langle w^2 \rangle = k \sum_{i=1}^{N} \langle (z_i - a)^2 \rangle .$$

Now, by definition, $\langle (z_i - \mu)^2 \rangle = \sigma^2$, from which we have

$$\langle w^2 \rangle = kN\sigma^2 \implies k = \frac{1}{N}.$$

Lesson: if a sample's population mean μ is known, we can construct an unbiased estimator of the population variance σ^2 according to

$$w^2 = \frac{1}{N} \sum_{i=1}^{N} (z_i - \mu)^2.$$

However, this is not a realistic situation. Normally the population mean μ is unknown, and we are limited to working with the sample mean $\overline{z}$, in which case we use the population variance estimator

$$s^2 = \frac{1}{N-1} \sum_{i=1}^{N} (z_i - \overline{z})^2 .$$

13.4 Theory: The Chi-Square Distribution

- The chi-square distribution is used when estimating the population variance corresponding to a normally-distributed sample $\{x_i\}_{i=1}^{N}$. Assuming $x_i \sim \mathcal{N}(0,1)$, we first construct the chi-square statistic

$$\chi^2 = \sum_{i=1}^{N} x_i^2.$$

Without derivation, the thus-constructed statistic χ^2 is distributed according to

$$\chi^2 \sim \frac{\mathrm{d}P}{\mathrm{d}\chi^2} \equiv \chi^2(\nu),$$

where $\chi^2(\nu)$ denotes the chi-square distribution with ν degrees of freedom. The distribution's degrees of freedom depend on the size of the sample; if $x_i \sim \mathcal{N}(0,1)$ as above then $\nu = N$.

Note that, confusingly, both the sample statistic χ^2 and the corresponding probability distribution function are denoted by χ^2.

- Now consider a general normally-distributed sample $\{z_i\}_{i=1}^{N}$ where $z_i \sim \mathcal{N}(\mu, \sigma^2)$. First, we construct the standardized variable

$$x_i = \frac{z_i - \mu}{\sigma} \sim \mathcal{N}(0,1),$$

and then construct the χ^2 sample statistic according to

$$\chi^2 = \sum_{i=1}^{N} \frac{(z_i - \mu)^2}{\sigma^2} \sim \chi^2(N), \tag{13.3}$$

which is distributed as a chi-square distribution with N degrees of freedom.

In Practice

- In practice a sample's population mean μ is unknown, so the construction of the chi-square statistic in Equation 13.3 is not useful. Instead, we replace the population mean μ with its estimator $\overline{z}$ to get

$$\chi^2 = \sum_{i=1}^{N} \frac{(z_i - \overline{z})^2}{\sigma^2} \sim \chi^2(N-1).$$

  Note that when constructed with $\overline{z}$ instead of μ, the chi-square statistic is distributed according to a chi-square distribution $\chi^2(N-1)$ with $N-1$ (and not N) degrees of freedom.

- Next, recalling the variance estimator s^2 in Equation 13.2, we can write the chi-square statistic as

$$\chi^2 = \sum_{i=1}^{N} \frac{(z_i - \overline{z})^2}{\sigma^2} = (N-1)\frac{s^2}{\sigma^2} \implies \sigma^2 = (N-1)\frac{s^2}{\chi^2}.$$

  Now, we know that the sample statistic χ^2 is distributed according to a chi-square distribution with $N-1$ degrees of freedom. The chi-square distribution is defined for $\chi^2 \in (0, \infty)$, so the sample statistic χ^2, and in turn σ^2, could fall anywhere in the range $(0, \infty)$. In practice, we find a confidence interval (σ^2_-, σ^2_+) for which we can expect, at a given confidence level, that the true population variance σ^2 falls in the range (σ^2_-, σ^2_+).

Confidence Interval

- Recall the equation

$$\sigma^2 = (N-1)\frac{s^2}{\chi^2}.$$

  Define upper and lower confidence bounds χ^2_+ and χ^2_- and corresponding bounds σ^2_+ and σ^2_-.

- The confidence interval (σ^2_-, σ^2_+) is associated with a *significance level* α. The probability that $\sigma^2 < \sigma^2_-$ is $\alpha/2$ and likewise the probability that $\sigma^2 > \sigma^2_+$ is $\alpha/2$. The probability that the true population variance σ^2 falls outside of the confidence interval (σ^2_-, σ^2_+) is then α.

- We find the upper confidence level χ^2_+ from the value of χ^2 satisfying

$$P(\chi^2 > \chi^2_+) = \frac{\alpha}{2}$$

  Find lower confidence limit χ^2_- from the value of χ^2 satisfying

$$P(\chi^2 < \chi^2_-) = \frac{\alpha}{2} \quad \text{or} \quad P(\chi^2 > \chi^2_-) = 1 - \frac{\alpha}{2}.$$

  Using the confidence levels χ^2_- and χ^2_+, we find the population variance confidence levels σ^2_+ and σ^2_- from

$$\sigma^2_\pm = (N-1)\frac{s^2}{\chi^2_\mp}.$$

  Note the signs of $\sigma^2_\pm$ and $\chi^2_\mp$ are reversed.

13.5 Student's t Distribution

- Student's t distribution is used when estimating the population mean associated with a normally-distributed sample $\{z_i\}_{i=1}^N \sim \mathcal{N}(\mu, \sigma^2)$. Using the sample values z_i, we first construct the sample statistic

$$T \equiv \frac{\overline{z} - \mu}{s}\sqrt{N}$$

where $\overline{z}$, s and N are sample's mean, variance, and size, respectively. Without proof, the T statistic is distributed according to

$$T \sim \frac{\mathrm{d}P}{\mathrm{d}T} = S(\nu),$$

where $S(\nu)$ is Student's t distribution with ν degrees of freedom; $\nu = N - 1$ for a T statistic constructed from a sample of size N.

- Next note that in terms of T, the population mean μ reads

$$\mu = \overline{z} - \frac{Ts}{\sqrt{N}}. \tag{13.4}$$

- Student's t distribution is symmetric, which motivates the definition of upper and lower confidence levels T_+ and T_- centered about $T = 0$, such that $T_- = -T_+$. We then use T_- and T_+ to construct corresponding confidence intervals $\mu_\pm$ for the population mean using Equation 13.4.

  A set of confidence levels μ_- and μ_+ are associated with a significance level α such that the probability that the population mean μ exceeds μ_+ is $\alpha/2$, while the probability that μ is less than μ_- is $\alpha/2$.

- We find the upper level T_+ from the value of T satisfying

$$P(|T| > T_+) = \alpha,$$

while the lower level T_- is trivially found from $T_- = -T_+$. With $T_\pm$ known, the corresponding confidence levels $\mu_\pm$ for the population mean are

$$\mu_\pm = \overline{z} - T_\pm \frac{s}{\sqrt{N}}.$$

13.6 Example: Simple Estimation of Population Mean

Given a sample

$$Z = \{1.162, 0.25, -0.052, 0.014, -0.362, 0.756, 1.62, 0.608, -2.15\},$$

which we may assume is drawn from a normally-distributed population $\mathcal{N}(\mu, \sigma^2)$, find the confidence interval $\mu_\pm$ for the population mean μ at the significance level $\alpha = 0.1$.

- We aim to estimate the population mean, so we will use the Student statistic

$$T = \frac{\overline{z} - \mu}{s}\sqrt{N}.$$

We first compute the sample mean $\overline{z}$ and variance s^2; these come out to

$$\overline{z} = \frac{1}{N}\sum_i z_i = 0.205 \qquad \text{and} \qquad s^2 = \frac{1}{N-1}\sum_i (z_i - \overline{z})^2 = (1.082)^2$$

The corresponding T statistic will be distributed according to Student's t distribution with $\nu = N - 1 = 8$ degrees of freedom.

- We find the upper confidence interval T_+ for the T statistic from the value of T_+ solving the equation $P(|T| > T_+)\alpha$; for $\nu = 8$ and $\alpha = 0.1$ this comes out to

$$T_+ = 1.860 \implies T_- = -T_+ = -1.860.$$

The corresponding confidence interval for μ, at the significance level $\alpha = 0.1$, is

$$\mu_\pm = 0.205 \pm 1.860 \cdot \frac{1.082}{3} \implies (\mu_-, \mu_+) = (0.877, -0.466).$$

14 Fourteenth Exercise Set

14.1 Theory: Hypotheses Testing

14.1.1 Mean Testing

- Consider a normally-distributed sample $\{z_i\}_{i=1}^N$ where $z_i \sim \mathcal{N}(\mu, \sigma^2)$. We then define a *null hypotheses* that $\mu = \mu_0$.
- Since we are working with the population mean μ we use the T statistic, which we construct under the hypotheses $\mu = \mu_0$ according to

  $$T_0 = \frac{\overline{z} - \mu_0}{s^2}\sqrt{N}$$

  This produces a value of T_0 under the hypotheses $\mu = \mu_0$.
- To test the null hypotheses $\mu = \mu_0$ at the significance level α, we first find the critical value T_c solving the equation

  $$P\big(|T| > T_\mathrm{c}\big) = \alpha,$$

  assuming T is distributed according to Student's t distribution with $\nu = N - 1$ degrees of freedom. Finally, we compare T_c to the value of T_0 computed under the hypotheses $\mu = \mu_0$. If $|T_0| > T_\mathrm{c}$, we reject the null hypotheses $\mu = \mu_0$ at the significance level α. We cannot, however, *confirm* the null hypotheses if $|T_0| < T_\mathrm{c}$.

14.1.2 Variance Testing

- In this case define the null hypotheses $\sigma^2 = \sigma_0^2$. We then compute the value of the chi-square sample statistic under the hypotheses $\sigma^2 = \sigma_0^2$ to get the value

  $$\chi_0^2 = (N-1)\frac{s^2}{\sigma^2}.$$

- To test the hypotheses $\sigma^2 = \sigma_0^2$ at the significance level α, we then find the values χ_+^2 and χ_-^2 for which

  $$P\big(\chi^2 > \chi_+^2\big) = \frac{\alpha}{2} \qquad \text{and} \qquad P\big(\chi^2 > \chi_-^2\big) = 1 - \frac{\alpha}{2},$$

  assuming that the sample statistic χ^2 is distributed according to a chi-square distribution with $\nu = N - 1$ degrees of freedom. Then if $\chi_0^2 < \chi_-^2$ or $\chi_0^2 > \chi_+^2$ (meaning that the chi-square value χ_0^2 found under the hypotheses $\sigma^2 = \sigma_0^2$ falls outside the confidence interval χ_-^2, χ_+^2), we reject the null hypotheses $\sigma^2 = \sigma_0^2$ at the significance level α.

14.1.3 Comparing Two Sample Means

- Consider two samples

  $$\{X_i\} : X_i \sim \mathcal{N}(\mu_x, \sigma^2) \qquad \text{and} \qquad \{Y_i\} : Y_i \sim \mathcal{N}(\mu_y, \sigma^2),$$

  which could represent, say, measurements of the same physical quantity with the same instrument but measured at two different times of the day. Importantly, we assume both samples have the same variance σ^2, although the mean values μ_x and μ_y may differ.

- We then use these samples to define the sample statistic

$$T = \frac{(\overline{x} - \overline{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{1}{N_x} - \frac{1}{N_y}}} \cdot \frac{\sqrt{N_x + N_y - 2}}{\sqrt{s_x^2(N_x - 1) + s_y^2(N_y - 1)}}, \tag{14.1}$$

  which, without proof, is distributed according to Student's t distribution with $\nu = N_x - N_y - 2$ degrees of freedom.

- Next, we define the equal-variance hypotheses $\mu_x = \mu_y$, which we use to compute the statistic in Equation 14.1. In this case $(\mu_x - \mu_y) = 0$, and we have

$$T_0 = \frac{(\overline{x} - \overline{y})}{\sqrt{\frac{1}{N_x} - \frac{1}{N_y}}} \cdot \frac{\sqrt{N_x + N_y - 2}}{\sqrt{s_x^2(N_x - 1) + s_y^2(N_y - 1)}}.$$

- To test the hypotheses $\mu_x = \mu_y$ at the significance level α, we compute the value of T_c solving the equation

$$P\big(|T| > T_\mathrm{c}\big) = \alpha,$$

  where we assume T is distributed according to Student's t distribution with $\nu = N_x - N_y - 2$ degrees of freedom.

  Finally, if $|T_0| > T_\mathrm{c}$, we reject the hypotheses $\mu_x = \mu_y$ at the significance level α.

14.2 Example: Hypotheses Testing of Sample Mean

14.2.1 Single-Sample Test

Given a six-measurement sample

$$x_i = \big\{0.44, 0.46, 050, 0.47, 0.48, 0.47\big\},$$

which we may assume is drawn from a normally-distributed population $\mathcal{N}(\mu, \sigma^2)$*, test the hypotheses* $\mu = 0.50 \equiv \mu_0$ *at the significance level* $\alpha = 0.1$*.*

- Since we are working with population mean, we will use the T statistic

$$T = \frac{\overline{x} - \mu}{s}\sqrt{N},$$

  which will distributed according to Student's t distribution with $\nu = N - 1 = 5$ degrees of freedom. We first make the intermediate calculations

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{6} x_i \approx 0.47 \qquad \text{and} \qquad s^2 = \frac{1}{N-1}\sum_{i=1}^{6}(x_i - \overline{x})^2 \approx (0.02)^2,$$

  and then compute the value of T under the hypotheses $\mu = \mu_0$. Using the just-computed values of $\overline{x}$ and s^2, this comes out to

$$T_0 = \frac{\overline{x} - \mu_0}{s}\sqrt{N} = -3.67.$$

- To test the hypotheses $\mu = \mu_0$ at the significance level α, we then find the critical T statistic value T_c solving the equation $P\big(|T| > T_\mathrm{c}\big) = \alpha$ for $\nu = 5$ degrees of freedom. With the help of a table, this comes out to

$$T_\mathrm{c}(\alpha = 0.1, k - 5) = 2.015.$$

- Finally, we test the hypotheses and find

$$|T_0| = 2.015 > T_{\rm c} = -3.67,$$

on the basis of which we reject the null hypotheses $\mu = \mu_0$ at the significance level $\alpha = 0.1$ for our given sample. Interpretation: the result $|T_0| > T_{\rm c}$ means that the value T_0 constructed under the hypotheses $\mu = \mu_0$ lies outside the confidence interval $[T_-, T_+]$ for the sample mean μ.

14.2.2 Two-Sample Test

We repeat the experiment above and get the new sample

$$y_i = \big\{0.52, 0.50, 0.55, 0.53, 0.52, 0.53\big\},$$

where we may assume $y_i \sim \mathcal{N}(\mu_x, \sigma^2)$. *Test the null hypotheses* $\mu_x = \mu_y$ *at the significance level* $\alpha = 0.1$.

- We first compute the sample mean and variance

$$\overline{y} \approx \frac{1}{N_y}\sum_{i=1}^{6} y_i = 0.525 \qquad \text{and} \qquad s_y^2 = \frac{1}{N_y - 1}\sum_{i=1}^{6}(y_i - \overline{y})^2 \approx (0.016)^2.$$

For review, the corresponding sample statistics for x sample were

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{6} x_i \approx 0.47 \qquad \text{and} \qquad s^2 = \frac{1}{N - 1}\sum_{i=1}^{6}(x_i - \overline{x})^2 \approx (0.02)^2.$$

- We then define the two-sample T statistic

$$T = \frac{(\overline{x} - \overline{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{1}{N_x} - \frac{1}{N_y}}} \cdot \frac{\sqrt{N_x + N_y - 2}}{\sqrt{s_x^2(N_x - 1) + s_y^2(N_y - 1)}},$$

which will be distributed according to the Student distribution with $\nu = N_x - N_y - 2$ degrees of freedom. Using this statistic, we then compute the value T_0 under the null hypotheses $\mu_x = \mu_y$. In this case $(\mu_x - \mu_y) = 0$ and we have

$$T_0 = \frac{(\overline{x} - \overline{y})}{\sqrt{\frac{1}{N_x} - \frac{1}{N_y}}} \cdot \frac{\sqrt{N_x + N_y - 2}}{\sqrt{s_x^2(N_x - 1) + s_y^2(N_y - 1)}} \approx 5.3$$

- To test the hypotheses $\mu_x = \mu_y$ at the significance level $\alpha = 0.1$, we then find the value $T_{\rm c}$ satisfying the equation

$$P\big(|T| > T_{\rm c}\big) = \alpha.$$

Using a table with $\nu = 10$, this comes out to

$$T_{\rm c}(\alpha = 0.1,\, \nu = 10) \approx 1.81.$$

We then perform the hypotheses test and find

$$|T_0| = 5.3 > T_{\rm c} = 1.81,$$

on the basis of which we reject the hypotheses $\mu_x = \mu_y$ at the significance level $\alpha = 0.1$ for our given samples $\{x_i\}$ and $\{y_i\}$.

14.3 Theory: Distribution Shape Tests

The goal of distribution shape tests (I don't known the conventional name) is to determine if a sample is distributed according to a given probability distribution. In symbols, given a sample $\{x_i\}$, we aim to test if $x_i \sim \frac{\mathrm{d}P}{\mathrm{d}x} = f(x)$ for a given probability distribution $f(x)$.

14.3.1 Pearson Chi-Square Test

- We first consider the Pearson chi-square test, in which case we denote chi by χ^2_{P}.

  Given a sample of N measurements $\{x_i\}$, draw a histogram of the sample values using an "appropriate number of bins" (clarified below) and count the number of measurements falling into each bin, where we let N_k denote the number of measurements falling into the k-th bin.

- We then compare each N_k to the predicted number of samples in the k-th bin based on the assumed probability distribution; this is

  $$N_{\mathrm{k}}^{\mathrm{pred}} = Np_{\mathrm{k}},$$

  where we of course require $N = \sum_k N_k$; we find the probability p_k from

  $$p_k = \int_{x_{k-1}}^{x_k} \frac{\mathrm{d}P}{\mathrm{d}x}\,\mathrm{d}x.$$

- Using the values of N_k and p_k and N we then construct the sample statistic

  $$\chi^2_{\mathrm{P}} = \sum_{k=1}^{B} \frac{(N_k - Np_{\mathrm{k}})^2}{Np_{\mathrm{k}}},$$

  where B is the number of bins. It turns out that the thus-constructed statistic χ^2_{P} is (approximately) distributed as

  $$\chi^2_{\mathrm{P}} \sim \chi^2(B-1),$$

  i.e. a chi-square distribution with $\nu = B - 1$ degrees of freedom. However, for $\chi^2_{\mathrm{P}} \sim \chi^2(B-1)$ to hold, we require $Np_k \gtrsim 5$ for each bin. If the condition $Np_k \gtrsim 5$ does not hold, we typically join adjacent bins to reach the required number Np_k of samples per bin.

- Assuming the condition $Np_k \gtrsim 5$ holds, we then aim to test the hypotheses that the sample $\{x_i\}$ is distributed according to $x_i \sim \frac{\mathrm{d}P}{\mathrm{d}x}$ at the significance level α. To do so, we first find the value of χ^2_{c} satisfying

  $$P(\chi^2 > \chi^2_c) = \alpha$$

  using a chi-square distribution with $B-1$ degrees of freedom. Finally, if $\chi^2_{\mathrm{P}} > \chi^2_{\mathrm{c}}$, we reject the assumed distribution $\frac{\mathrm{d}P}{\mathrm{d}x}$ for the $\{x_i\}$ at the significance level α.

14.4 Example: Pearson Chi-Square Test

We roll a six-side die 60 *times and get the results*

Outcome	Number of events
1	5
2	8
3	9
4	8
5	10
6	20

Use the Pearson chi-square test to determine if the die is fair at the two confidence levels 0.99 and 0.95, i.e. $\alpha = 0.01$ *and* $\alpha = 0.05$.

- We divide the sample into the obvious choice of 6 bins, one for each possible outcome. In this case the N_k are just the number of events for each outcome in the table above, and $N = \sum_k N_k = 60$.
- Mathematically, a fair die means the outcomes are distributed according to a uniform distribution, i.e. $p_k = 1/6$ for $k = 1, 2, \ldots, 6$. Assuming a uniform distribution, we then find that

$$Np_k = 60 \cdot \frac{1}{6} = 10 \text{ for } k = 1, \ldots, 6.$$

  Since $Np_k > 5$ for all k, we can use the Pearson chi-square test.

- We then construct the Pearson chi-square statistic according to

$$\chi^2_{\mathrm{P}} = \sum_{k=1}^{B} \frac{(N_k - Np_{\mathrm{k}})^2}{Np_{\mathrm{k}}},$$

  Substituting in the values of N_k, p_k and N, we find

$$\begin{aligned}\chi^2_{\mathrm{P}} &= \frac{(5-10)^2}{10} + \frac{(8-10)^2}{10} + \frac{(9-10)^2}{10} + \frac{(8-10)^2}{10} + \frac{(10-10)^2}{10} + \frac{(20-10)^2}{10} \\ &= 2.5 + 0.4 + 0.1 + 0.4 + 0.0 + 10 \\ &= 13.4.\end{aligned}$$

- First, for $\alpha = 0.01$ and using a chi-square distribution with $\nu = B - 1 = 5$ degrees of freedom, we use a table of values to solve the equation

$$P(\chi^2 > \chi^2_{\mathrm{c}}) = \alpha \implies \chi^2_{\mathrm{c}} = 15.086.$$

  We then perform the Pearson chi-square test and find

$$\chi^2_{\mathrm{P}} = 13.4 < \chi^2_{\mathrm{c}} = 15.086,$$

  which means we *cannot* reject the hypotheses that the dies' outcomes are uniformly distributed at the significance level $\alpha = 0.99$.

- Meanwhile, for $\alpha = 0.05$, we find

$$P(\chi^2 > \chi^2_c) = \alpha \implies \chi^2_c = 11.07.$$

In this case we perform the Pearson chi-square test and find

$$\chi^2_P = 13.4 > \chi^2_c = 11.07,$$

which means we *can* reject the hypotheses that the dice is uniformly distributed (i.e. that the dice is fair) at the significance level $\alpha = 0.05$.

14.5 Theory: Linear Least Squares

- Given vector of N measurements $\boldsymbol{z} \in \mathbb{R}^N$, which each measurement z_i has the same uncertainty σ^2, we aim to find the optimal values of the M parameters $\boldsymbol{x} \in \mathbb{R}^M$ for the model

$$\boldsymbol{z} = \mathbf{H}\boldsymbol{x} + \boldsymbol{r}. \tag{14.2}$$

Concretely, we aim to find the parameter values $\widehat{\boldsymbol{x}}$ minimizing the quadratic form

$$\chi^2 = \boldsymbol{r}^\top \mathbf{R}^{-1} \boldsymbol{r}, \quad \text{where } \mathbf{R} = \sigma^2 \mathbf{I}.$$

Solving for $\boldsymbol{r}$ in Equation 14.2, we see that χ^2 is computed according to

$$\chi^2 = \big(\boldsymbol{z} - \mathbf{H}\boldsymbol{x}\big)^\top \mathbf{R}^{-1} (\boldsymbol{z} - \mathbf{H}\boldsymbol{x}).$$

Importantly, the quadratic form χ^2 is distributed as $\chi^2 \sim \chi^2(N - M)$, i.e. a chi-square distribution with $\nu = N - M$ degrees of freedom.

- Without derivation, the process of minimizing χ^2 with respect to the model parameters $\boldsymbol{x}$ results in the equations

$$\widehat{\boldsymbol{x}} = (\mathbf{H}^\top \mathbf{H})^{-1} \mathbf{H}^\top \boldsymbol{z} \qquad \text{and} \qquad \mathbf{P} = (\mathbf{H}^\top \mathbf{H})^{-1} \sigma^2,$$

where $\widehat{\boldsymbol{x}}$ holds the optimal parameter values and $\mathbf{P}$ is the corresponding covariance matrix for the optimal parameters $\widehat{\boldsymbol{x}}$. **TODO:** interpret covariance matrix.

- To find the *structure matrix* $\mathbf{H}$, we first define a model

$$z_i = x_1 f_1(t_i) + \cdots x_M f_M(t_i) + r_i \text{ for } i = 1, 2, \ldots, N,$$

where M is the number of parameters. The structure matrix $\mathbf{H}$ is then

$$\mathbf{H} = \begin{pmatrix} f_1(t_1) & f_1(t_1) & \cdots & f_M(t_1) \\ f_1(t_2) & f_1(t_2) & \cdots & f_M(t_2) \\ \vdots & \cdots & \ddots & \vdots \\ f_1(t_N) & f_1(t_N) & \cdots & f_M(t_N) \end{pmatrix}$$

Measurements with Different Variances

- If each measurement z_i has a different variance, e.g. σ_i^2 instead of σ^2, we then have

$$(\mathbf{R})_{ij} = \sigma_i^2 \delta_{ij}$$

In this case the quadratic form χ^2 generalizes to

$$\chi^2 = \big(\boldsymbol{z} - \mathbf{H}\boldsymbol{x}\big)^\top \mathbf{R}^{-1} (\boldsymbol{z} - \mathbf{H}\boldsymbol{x}),$$

while the optimal estimate and covariance matrix become

$$\widehat{\boldsymbol{x}} = (\mathbf{H}^\top \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^\top \mathbf{R}^{-1} \boldsymbol{z} \qquad \text{and} \qquad \mathbf{P} = (\mathbf{H}^\top \mathbf{R}^{-1} \mathbf{H})^{-1}.$$

15 Fifteenth Exercise Set

15.1 Example: Linear Least Squares for $z_i = x_0 t_i + r_i$

- Consider a model with a single parameter x_0 of the form

$$z_i = x_0 t_i + r_i.$$

We will first determine structure matrix $\mathbf{H}$. We denote a hypothetical vector of N measurements as

$$\boldsymbol{z} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{pmatrix},$$

while the vector of parameters for our simple one-parameter model is just $\boldsymbol{x} = (x_0)$. Recalling the general expression $z = x_0 f_0(t_i) + r_i$, we see that in our case $f_0 = t_i$, and the structure matrix reads

$$\mathbf{H} = \begin{pmatrix} f_0(t_1) \\ f_0(t_2) \\ \vdots \\ f_0(t_N) \end{pmatrix} = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix}.$$

- Next we compute

$$\mathbf{H}^\top \mathbf{H} = \begin{pmatrix} t_1 \, t_2 \, \cdots t_N \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix} = \sum_{i=1}^{N} t_i^2 \qquad \text{and} \qquad \mathbf{H}^\top \boldsymbol{z} = \begin{pmatrix} t_1 \, t_2 \, \cdots t_N \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} = \sum_{i=1}^{N} z_i t_i.$$

From here we find the covariance "matrix"

$$\mathbf{P} = (\mathbf{H}^\top \mathbf{H})^{-1} \sigma^2 = \frac{\sigma^2}{\sum_{i=1}^{N} t_i^2} = \sigma_{\widehat{x}_0}^2,$$

while the corresponding optimal parameter is

$$\widehat{\boldsymbol{x}} = (\mathbf{H}^\top \mathbf{H})^{-1} \mathbf{H}^\top \boldsymbol{z} = \frac{\sum_{i=1}^{N} t_i z_i}{\sum_{i=1}^{N} t_i^2} = \widehat{x}_0.$$

15.2 Theory: Linear Least Squares for $z_i = x_0 t_i + x_1 + r_i$

- In this section we will consider in detail the two-parameter model

$$z_i = x_0 t_i + x_1 + r_i.$$

The model's parameter vector and structure matrix are

$$\boldsymbol{x} = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \qquad \text{and} \qquad \mathbf{H} = \begin{pmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & 1 \\ t_N & 1 \end{pmatrix}.$$

- First, we compute the product

$$\mathbf{H}^\top \mathbf{H} = \begin{pmatrix} t_1 & t_2 & \cdots & t_N \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & 1 \\ t_N & 1 \end{pmatrix} = \begin{pmatrix} \sum t_i^2 & \sum t_i \\ \sum t_i & N \end{pmatrix},$$

  where the sums runs over i from 1 to N. We then compute inverse of $\mathbf{H}^\top \mathbf{H}$, using the general 2×2 matrix identity

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \implies \mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

  Applied to $\mathbf{H}^\top \mathbf{H}$ we have

$$\left(\mathbf{H}^\top \mathbf{H}\right)^{-1} = \frac{1}{N \sum t_i^2 - \left(\sum t_i\right)^2} \begin{pmatrix} N & -\sum t_i \\ -\sum t_i & \sum t_i^2. \end{pmatrix}$$

  The model's corresponding covariance matrix is then $\mathbf{P} = \left(\mathbf{H}^\top \mathbf{H}\right)^{-1} \sigma^2$, assuming the measurements $\{z_i\}$ used to construct the model all have the same variance σ^2.

- To find the optimal parameters $\widehat{\boldsymbol{x}}$, we first max the auxiliary calculation

$$\mathbf{H}^\top \boldsymbol{z} = \begin{pmatrix} t_1 & t_2 & \cdots & t_N \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} \sum t_i z_i \\ \sum z_i \end{pmatrix},$$

  in terms of which the optimal parameters $\widehat{x}$ are then

$$\widehat{\boldsymbol{x}} = \left(\mathbf{H}^\top \mathbf{H}\right)^{-1} \mathbf{H}^\top \boldsymbol{z} = \frac{1}{N \sum t_i^2 - \left(\sum t_i\right)^2} \begin{pmatrix} N \sum t_i z_i - \left(\sum t_i\right) \cdot \left(\sum z_i\right) \\ -\left(\sum t_i\right) \cdot \left(\sum t_i z_i\right) + \left(\sum t_i^2\right) \cdot \left(z_i\right) \end{pmatrix}.$$

- Next, for shorthand, we define the constant

$$C = N \sum t_i^2 - \left(\sum t_i\right)^2,$$

  in terms of which the individual parameters $\widehat{x}_0$ and $\widehat{x}_1$ read

$$\widehat{x}_0 = \frac{1}{C} \left[N \sum t_i z_i - \left(\sum t_i\right) \cdot \left(\sum z_i\right) \right], \tag{15.1}$$

$$\widehat{x}_1 = \frac{1}{C} \left[\left(\sum t_i^2\right) \cdot \left(\sum z_i\right) - \left(\sum t_i\right) \cdot \left(\sum t_i z_i\right) \right] \tag{15.2}$$

- Next, we aim to express x_1 in terms of x_0. We first multiply the equation for $\widehat{x}_0$ through by $C \sum t_i$ to get

$$C \widehat{x}_0 \left(\sum t_i\right) = N \left(\sum t_i\right) \cdot \left(\sum t_i z_i\right) - \left(\sum t_i\right)^2 \left(\sum z_i\right),$$

  and then multiply the equation for $\widehat{x}_1$ through by CN to get

$$C \widehat{x}_1 N = N \left(\sum t_i^2\right) \left(\sum z_i\right) - N \left(\sum t_i\right) \left(\sum t_i z_i\right).$$

We then add these equations and cancel common terms to get

$$\begin{aligned} C\left[\widehat{x}_0\left(\sum t_i\right)+\widehat{x}_1 N\right] &= N\left(\sum t_i^2\right)\left(\sum z_i\right)-\left(\sum t_i\right)^2\left(\sum z_i\right) \\ &= \left(\sum z_i\right)\left[N\sum t_i^2-\left(\sum t_i\right)^2\right] \\ &= \left(\sum z_i\right)\cdot C. \end{aligned}$$

We then cancel C from both sides of the equation to get

$$\widehat{x}_0\sum t_i+\widehat{x}_1 N=\sum z_i,$$

from which we can express $\widehat{x}_1$ in terms of $\widehat{x}_0$ as

$$\widehat{x}_1=\frac{1}{N}\sum z_i-\frac{1}{N}\left(\sum t_i\right)\widehat{x}_0.$$

But the first is just the average of the measurements $\overline{z}$ and second term's coefficient is the average $\overline{t}$ of the independent variable. We thus have

$$\widehat{x}_1=\overline{z}-\overline{t}\widehat{x}_0.$$

Next: $\widehat{x}_0$

- Recall from Equation 15.1 that $\widehat{x}_0$ is given by

$$\begin{aligned} \widehat{x}_0 &= \frac{1}{N\sum t_i^2+\left(\sum t_i\right)^2}\cdot\left[N\sum t_i z_i-\sum t_i\sum z_i\right]. \\ &= \frac{1}{N\sum t_i^2+-N^2\overline{t}^2}\cdot\left[N\sum t_i z_i-N^2\overline{t}\overline{z}\right], \end{aligned}$$

where we have recognized expressions for $N\overline{t}$ and $N\overline{z}$ in square bracket term and $\overline{t}$ in denominator. We then cancel common N terms, factor out remaining N, then recognize $\overline{t^2}$ in denominator to get

$$\widehat{x}_0=\frac{\overline{tz}-\overline{t}\overline{z}}{\overline{t^2}-\overline{t}^2}$$

Lesson: an equivalent expression for finding $\overline{x}_0$ for the linear model $z_i=x_0 t_i+x_1+r_i$. We then can find $\widehat{x}_1$ from $\widehat{x}_1=\overline{z}-\overline{t}\widehat{x}_0$.

- Main takeaway is to have equations for optimal parameters ready instead of finding these by definition from structure matrix.

15.3 Exercise: Linear Least Squares with a Quadratic Model

We are given data for a car's fuel consumption C as a function of its speed v:

v [km h^{-1}]	C [L/100 km]
60	4.8
72	5.0
90	7.1
120	8.2
150	11.0

The uncertainty on each measurement is $\sigma_C^2 = 0.5\,\mathrm{L}/100\,\mathrm{km}$. *At the significance level* $\alpha = 0.1$, *test the hypotheses that the consumption obeys the quadratic model* $C = C_0 + \beta v^2$.

First Part

- We will first find the optimal parameter values $\widehat{C}_0$ and $\widehat{\beta}$, then find the corresponding form χ^2.
- To match the model $C = C_0 + \beta v^2$ to the model $z_i = x_0 t_i + x_1$ from the previous section, we first introduce the new notation $v^2 \to t$, $C \to z$ and

$$\boldsymbol{x} = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \to \begin{pmatrix} \beta \\ C_0 \end{pmatrix}.$$

We can then directly use the previous section's results, which reveal that the optimal model parameters are

$$\widehat{x}_0 = \widehat{\beta} = \frac{\overline{tz} - \bar{t}\bar{z}}{\overline{t^2} - \bar{t}^2} \qquad \text{and} \qquad \widehat{x}_1 = \widehat{C}_0 = \bar{z} - \bar{t}\widehat{x}_0. \tag{15.3}$$

To find numeric values, we have to first compute $\bar{z}$, $\bar{t}$, $\overline{t^2}$ and $\overline{tz}$. Using a table for better organization, we have

$z = C_0$	v	$t = v^2$	$t^2\,[\times 10^4]$	tz
4.8	60	3600	1296	17280
5.0	72	5184	2687	25920
7.1	90	8100	6561	57510
8.2	120	14400	20736	118080
11.0	150	22500	50625	247500

From the table we find $\bar{z} = 7.22$ and $\bar{t} = 10756.8$ and $\overline{t^2} = 16381 \cdot 10^4$ and $\overline{tz} = 93258$, which we substitute into Equation 15.3 to get

$$\widehat{\beta} = 3.24 \cdot 10^{-4}\,(\mathrm{L}/100\,\mathrm{km}) \cdot \mathrm{km\,h^{-2}} \qquad \text{and} \qquad \widehat{C}_0 = 3.73\,\mathrm{L\,km^{-1}}.$$

Part Two: Confidence Interval

- Begin with the general linear least squares quadratic form expression

$$\chi^2 = \left(\boldsymbol{z} - \mathbf{H}\widehat{\boldsymbol{x}}\right)^\top \mathbf{R}^{-1} (\boldsymbol{z} - \mathbf{H}\widehat{\boldsymbol{x}})$$

In our case the $\boldsymbol{z}$ is simply the vector of measurements C_i, $\mathbf{H}\widehat{\boldsymbol{x}}$ is the predicted consumption on the basis of the optimal model parameters, which we will denote by $C(v_i; \widehat{\boldsymbol{x}})$, and $\mathbf{R} = \sigma^2 \mathbf{I}$, in terms of which χ^2 comes out to

$$\chi^2 = \frac{1}{\sigma^2} \sum_{i=1}^{N} \Big[C_i - C(v_i; \widehat{\boldsymbol{x}}) \Big]^2 \tag{15.4}$$

We then find the values C_i and $C(v_i; \widehat{\boldsymbol{x}})$ for each measurement v_i. The results appear in the table below

C_i	$C(v_i, \widehat{\boldsymbol{x}})$
4.8	4.90
5.0	5.41
7.1	6.35
8.2	8.40
11.0	11.02

We then substitute these values into Equation 15.4 and find

$$\chi^2 = \frac{1}{\sigma^2} \sum_{i=1}^{N} \Big[C_i - C(v_i; \widehat{\boldsymbol{x}}) \Big]^2 = 3.1 \equiv \chi_0^2.$$

- Finally, we aim to test the hypotheses, at $\alpha = 0.1$, that the given data points are indeed distributed according to the quadratic model $C = C_0 + \beta v^2$.

  Recall that the chi-square statistic in linear least squares obeys a chi-square distribution with $\nu = N - M$ degrees of freedom (N being number of data points and M number of parameters). Our current problem has $N = 5$ and $M = 2$, so $\nu = 5 - 2 = 3$. We then find the one-sided confidence limit χ_c^2 solving the equation

$$P(\chi^2 > \chi_c^2) = \alpha \implies \chi_{c_+}^2(\alpha = 0.1, \nu = 3) = 6.25.$$

  Performing the hypothesis test, we find

$$\chi_0^2 = 3.1 < \chi_c^2 = 6.25,$$

  which means that we cannot reject the quadratic model on the basis of our measurements at the significance level $\alpha = 0.1$.

TODO: Note

- Situation: we trust our model fits our data, but we don't know the values σ^2 of the datapoints (we assume same σ for all data points). ITC we can find a confidence interval for the value of σ.
- For review we wrote

$$\chi^2 = \sum_i \frac{(P_i - P(v_i; \widehat{\boldsymbol{x}}))^2}{\sigma^2}$$

  We then find interval on σ^2 via

$$\sigma_\pm^2 = \frac{1}{\chi_\mp^2} \sum \Big[P_i - P(v_i; \widehat{\boldsymbol{x}}) \Big]^2$$

  TODO: determine how to determine limits $\chi_\mp^2$ probably like above but use a two-sided chi-square test?

- Note, in passing: reduced chi-squred χ_r^2. If the variable χ^2 is distributed as $\chi^2 \sim \chi^2(N - M)$, the expected value of χ^2 turns out to be the number of degrees of freedom, i.e. $N - M$.

  The reduced chi-square is $\chi_r^2 = \frac{\chi^2}{N-M}$ which implies, because of normalization, we have $\langle \chi_r^2 \rangle = 1$, and then a value of χ_r^2 near one means "good agreement".