

(2) (a) $f(x; c) = \frac{x^2 + c}{2x}$ for $x_n \mapsto x_{n+1} = f(x_n; c)$

Fixed point means $x_n \mapsto x_n$, i.e.,

$$x_n = \frac{x_n^2 + c}{2x_n} \Rightarrow 2x_n^2 = x_n^2 + c$$

$$x_n^2 = c \Rightarrow \boxed{x_n = \pm \sqrt{c}}$$

Locations suggest this might be used to compute square roots.

(b) Given $x_0 = \frac{1}{2}$ and $c = 4$

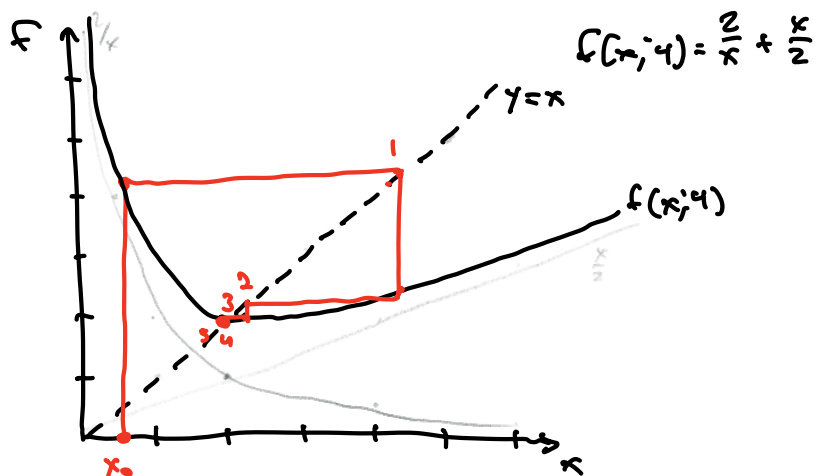
$$x_1 = \frac{(\frac{1}{2})^2 + 4}{2 \cdot \frac{1}{2}} = \frac{17}{4} = 4.25$$

$$x_2 = \frac{(\frac{17}{4})^2 + 4}{2 \cdot \frac{17}{4}} = \frac{\frac{289}{16} + 4}{\frac{17}{2}} = \frac{289 + 16 \cdot 4}{8 \cdot 17} = \frac{353}{136} \approx 2.59559\dots$$

$$x_3 = \frac{(\frac{353}{136})^2 + 4}{2 \cdot \frac{353}{136}} = \frac{124609 + 136^2 \cdot 4}{2 \cdot 136 \cdot 353} = \frac{198593}{96016} \approx 2.06833\dots$$

$$x_4 = \frac{(198593)^2 + (96016)^2 \cdot 4}{2 \cdot 198593 \cdot 96016} = \frac{76315468673}{38130210976} \approx 2.00113\dots$$

$$x_5 = \frac{(2.00113\dots)^2 + 4}{2 \cdot (2.00113\dots)} \approx 2.00000032\dots$$



Q1 cont'd (c) $f(x > 0; c) = \frac{x^2 + c}{2x} = \frac{\text{positive}}{\text{positive}}$ so $f > 0$ if $x > 0$.

Also $f(x; c) = \frac{x^2 + c}{2x} = -\frac{(-x)^2 + c}{2(-x)} = -f(-x; c)$, excluding $x=0$.
 Positive x_n iterates to positive x_{n+1} , and negative to negative.
 The iteration will approach the fixed point $x_* = -\sqrt{c}$.

(d) We have $x_n \mapsto f(x_n; c) = x_{n+1}$.

We want a map $w_n \mapsto w_{n+1}$ where $w = x - x_*$.

$$w_{n+1} = f(w_n + x_*; c) - x_* = \frac{(w_n + x_*)^2 + c}{2(w_n + x_*)} - x_*$$

$$= \frac{w_n^2 + 2w_n x_* + x_*^2 + c - 2x_*(w_n + x_*)}{2(w_n + x_*)} = \frac{w_n^2 - x_*^2 + c}{2(w_n + x_*)}$$

From (c), $x_*^2 = c$, so $x_* = \sqrt{c}$; $w_{n+1} = \frac{w_n^2}{2(w_n + \sqrt{c})}$

The mapping $x_n \mapsto x_{n+1}$ will get closer to x_* at iteration n

if $\left| \frac{w_{n+1}}{w_n} \right| = \left| \frac{w_n}{2(w_n + \sqrt{c})} \right| < 1$

We want $\left| \frac{w_0}{2(w_0 + \sqrt{c})} \right| < 1$, or $\frac{w_0^2}{4(w_0 + \sqrt{c})^2} < 1$, or

$w_0^2 < 4w_0^2 + 4w_0\sqrt{c} + 4c$, or $(w_0 + 2\sqrt{c})(3w_0 + 2\sqrt{c}) > 0$, which are true unless $-2\sqrt{c} < w_0 < -\frac{2}{3}\sqrt{c} \Rightarrow -\sqrt{c} < x_0 < \frac{1}{3}\sqrt{c}$.

We also need $x_0 > 0$ or iteration will converge to $-\sqrt{c}$.

Thus we need $x_0 > \frac{1}{3}\sqrt{c}$. (NB: Not satisfied in (b).)

If $0 < x_0 \leq \frac{1}{3}\sqrt{c}$, then $x_1 > \frac{1}{3}\sqrt{c}$ and convergence follows. So $x > 0$ is a basin of attraction for $x_* = \sqrt{c}$.

Q2 Given $\frac{dx}{dt} = y - 2b \sinh x$ (i)

$\frac{dy}{dt} = -x$. (ii)

(a) Differentiate (i) with respect to t , then insert (ii)

$$\frac{d^2x}{dt^2} = \frac{dy}{dt} - 2b \cosh x \frac{dx}{dt} \quad (\text{note chain rule})$$

$$\frac{d^2x}{dt^2} = -x - 2b \cosh x \frac{dx}{dt}$$

$$\Rightarrow \frac{d^2x}{dt^2} + 2b \cosh x \frac{dx}{dt} + x = 0.$$

This is the form $x'' + f(x)x' + x = 0$ with $f(x) = 2b \cosh x$

$f = 0$, simple harmonic oscillator; $f > 0$, damped.
 $f < 0$, driven.

(b) Equilibrium (x^*, y^*) of (i), (ii) when

$$0 = y^* - 2b \sinh x^* \quad \Rightarrow \quad (x^*, y^*) = (0, 0)$$

$$0 = -x^*$$

$$\underline{J}|_{(0,0)} = \begin{bmatrix} \frac{\partial (y - 2b \sinh x)}{\partial x} & \frac{\partial (y - 2b \sinh x)}{\partial y} \\ \frac{\partial (-x)}{\partial x} & \frac{\partial (-x)}{\partial y} \end{bmatrix}_{(0,0)} = \begin{bmatrix} -2b & 1 \\ -1 & 0 \end{bmatrix}$$

$$\det(\underline{J}|_{(0,0)} - \lambda \underline{I}) = (-2b - \lambda)(-\lambda) + 1 = \lambda^2 + 2b\lambda + 1 = 0$$

$$\lambda = -b \pm \sqrt{b^2 - 1} \quad \operatorname{Re}(\lambda) \neq 0 \text{ if } b \neq 0$$

$(0, 0)$ is a hyperbolic equilibrium point for all $b \neq 0$.

$$\textcircled{Q2} \textcircled{c} V(x, y) = x^2 + \frac{1}{2}y^2 + \frac{1}{2}(2b \sinh x - y)^2$$

$V(x, y)$ is suitable as a Lyapunov function because $V(0, 0) = 0$ and $V(x, y) > 0$ for all $(x, y) \neq (0, 0)$.

$$\frac{dV}{dt} = \frac{\partial V}{\partial x} \frac{dx}{dt} + \frac{\partial V}{\partial y} \frac{dy}{dt} \quad \textcircled{iii}$$

$$\frac{\partial V}{\partial x} = 2x + (2b \sinh x - y) \cdot 2b \cosh x \quad \textcircled{i}$$

$$\frac{\partial V}{\partial y} = y + (2b \sinh x - y)(-1) \quad \textcircled{iv}$$

Using \textcircled{i} , \textcircled{ii} , \textcircled{iv} , \textcircled{v} in \textcircled{iii}

$$\begin{aligned} \frac{dV}{dt} &= [2x + 2b \cosh x (2b \sinh x - y)](y - 2b \sinh x) \\ &\quad + [y - (2b \sinh x - y)](-x) \end{aligned}$$

$$= 2xy - 4bx \cosh x - 2b \cosh x (2b \sinh x - y)^2 - 2xy + 2bx \sinh x$$

$$= -2bx \sinh x - 2b \cosh x (2b \sinh x - y)^2$$

$$= -2b [x \sinh x + \cosh x (2b \sinh x - y)^2]$$

$$\frac{dV}{dt} \leq 0 \text{ for all } b > 0; \quad \frac{dV}{dt} = 0 \text{ only if } x = y = 0$$

Therefore the equilibrium point $(0, 0)$ is STABLE.

\textcircled{d} $V(x, y)$ has the property that $\lim_{\sqrt{x^2+y^2} \rightarrow \infty} V(x, y) = \infty$

So the origin is ASYMPTOTICALLY STABLE.

$$\ddot{y} + 2\varepsilon \dot{y} + y = 0 \quad y(0) = 0, \quad \dot{y}(0) = 1$$

(a) let $y = y_0 + \varepsilon y_1 + \dots$

$$(\ddot{y}_0 + \varepsilon \ddot{y}_1 + \dots) + 2\varepsilon(\dot{y}_0 + \varepsilon \dot{y}_1 + \dots) + (y_0 + \varepsilon y_1 + \dots) = 0$$

$$y_0(0) + \varepsilon y_1(0) + \dots = 0, \quad \dot{y}_0(0) + \varepsilon \dot{y}_1(0) + \dots = 1$$

$$\varepsilon^0: \left. \begin{array}{l} \ddot{y}_0 + y_0 = 0 \\ y_0(0) = 0, \dot{y}_0(0) = 1 \end{array} \right\} \left. \begin{array}{l} y_0 = A \sin(t) + B \cos(t) \\ y_0(0) = B = 0 \\ \dot{y}_0(0) = A = 1 \end{array} \right\} y_0(t) = \sin(t)$$

[1]
[1]

$$\varepsilon^1: \left. \begin{array}{l} \ddot{y}_1 + y_1 = -2\dot{y}_0 = -2\cos(t) \\ y_1(0) = \dot{y}_1(0) = 0 \end{array} \right\} y_1(t) = (C + Dt) \sin(t) + (E + Ft) \cos(t)$$

$$\dot{y}_1 = D \sin(t) + (C + Dt) \cos(t) + F \cos(t) - (E + Ft) \sin(t)$$

$$\begin{aligned} \ddot{y}_1 &= D \cos(t) + D \cos(t) - (C + Dt) \sin(t) - F \sin(t) - F \sin(t) - (E + Ft) \cos(t) \\ &= -(C + 2F + Dt) \sin(t) - (E - 2D + Ft) \cos(t) \end{aligned}$$

$$\ddot{y}_1 + y_1 = -2F \sin(t) + 2D \cos(t) = -2 \cos(t) \rightarrow F = 0, D = -1$$

$$y_1 = (C - t) \sin(t) + E \cos(t), \quad \dot{y}_1 = -(1 + E) \sin(t) + (C - t) \cos(t)$$

$$y_1(0) = E = 0, \quad \dot{y}_1(0) = C = 0 \rightarrow y_1(t) = -t \sin(t)$$

$$y(t) = \sin(t) - \varepsilon t \sin(t)$$

[1]

The second term is secular (grows w/o bound and will eventually disorder the series), so this solution is ok as long as $\varepsilon t \ll 1 \rightarrow t \ll 1/\varepsilon$

[1]

$$(b) \quad \lambda^2 + 2\varepsilon\lambda + 1 = 0 \rightarrow \lambda = \frac{-2\varepsilon \pm \sqrt{4\varepsilon^2 - 4}}{2} = -\varepsilon \pm i\sqrt{1-\varepsilon^2} \quad [1]$$

$$\rightarrow y = e^{-\varepsilon t} [A \sin(\alpha t) + B \cos(\alpha t)]$$

$$\dot{y} = e^{-\varepsilon t} [(-\varepsilon A - \alpha B) \sin(\alpha t) + (-\varepsilon B + \alpha A) \cos(\alpha t)]$$

$$y(0) = B = 0$$

$$\dot{y}(0) = -\varepsilon B + \alpha A = 1 \rightarrow A = \frac{1 + \varepsilon B}{\alpha} = \frac{1}{\alpha}$$

$$y(t) = \frac{1}{\sqrt{1-\varepsilon^2}} e^{-\varepsilon t} \sin(t\sqrt{1-\varepsilon^2}) \quad [1]$$

The exact solution decays exponentially in time, whereas the approximate one grows linearly in time. [1]

(c) Can see from the exact solution that we need to allow for a slow scale εt .

$$\rightarrow \text{let } t_1 = t, \quad t_2 = \varepsilon t \rightarrow \frac{d}{dt} = \frac{\partial t_1}{\partial t} \frac{\partial}{\partial t_1} + \frac{\partial t_2}{\partial t} \frac{\partial}{\partial t_2} = \frac{\partial}{\partial t_1} + \varepsilon \frac{\partial}{\partial t_2} \quad [1]$$

$$\left(\frac{\partial}{\partial t_1} + \varepsilon \frac{\partial}{\partial t_2}\right) \left(\frac{\partial}{\partial t_1} + \varepsilon \frac{\partial}{\partial t_2}\right) y + 2\varepsilon \left(\frac{\partial}{\partial t_1} + \varepsilon \frac{\partial}{\partial t_2}\right) y + y = 0 \quad [1]$$

$$y(0) = 0, \quad \frac{\partial y}{\partial t_1} \Big|_{0,0} + \varepsilon \frac{\partial y}{\partial t_2} \Big|_{0,0} = 1$$

$$\text{let } y = y_0 + \varepsilon y_1 + \dots$$

$$\varepsilon^0: \quad \frac{\partial^2 y_0}{\partial t_1^2} + y_0 = 0$$

$$y_0(0,0) = 0, \quad \frac{\partial y_0}{\partial t_1} \Big|_{0,0} = 1$$

$$y_0 = a(t_2) \sin(t_1) + b(t_2) \cos(t_1)$$

$$y_0(0,0) = 0: \quad b(0) = 0$$

$$\frac{\partial y_0}{\partial t_1} \Big|_{0,0} = 1: \quad a(0) = 1$$

$$\epsilon^1: \frac{\partial^2 y_1}{\partial t_1^2} + y_1 = -2 \frac{\partial y_0}{\partial t_1 \partial t_2} - 2 \frac{\partial y_0}{\partial t_1} = -2 \left[\underbrace{\left(a + \frac{da}{dt_2} \right) \cos(t_1)} - \underbrace{\left(b + \frac{db}{dt_2} \right) \sin(t_1)} \right]$$

must vanish to avoid secular terms in y_1

$$y(t) = e^{-\epsilon t} \sin(t)$$

$$\frac{da}{dt_2} + a = 0$$

$$a = C_1 e^{-t_2}$$

$$a(0) = 1 \implies C_1 = 1$$

$$a(t_2) = e^{-t_2}$$

$$\frac{db}{dt_2} + b = 0$$

$$b = C_2 e^{-t_2}$$

$$b(0) = 0 \implies C_2 = 0$$

$$b(t_2) = 0$$

Better b/c we now capture the exponential decay.

(2) Approximate solution still has slightly the wrong amplitude (1 vs. $\frac{1}{\sqrt{1-\epsilon^2}}$) and slightly the wrong frequency (1 vs. $\sqrt{1-\epsilon^2}$).

Strategies for improving:

- Go to higher order in ϵ - will gradually fix amplitude
- include more scales - will gradually fix period

1. Advanced Probability

- (a) Suppose the height above ground of the flight trajectory of an unidentified flying object (UFO) is given by the function $h : \mathbb{R} \rightarrow \mathbb{R}_+$ where \mathbb{R}_+ is the set of positive real numbers, $h(t)$ denotes the height (in m) at time t . Consider the Gaussian process $\mathcal{GP}(\mu, k)$ where the covariance function is given by $k(t, t') = \exp(-4|t - t'|^2)$ and where the mean function is $\mu(t) = 10, \forall t, t'$.

- (i) Argue briefly whether or not this Gaussian process might be a reasonable a priori model of a subjective (Bayesian) belief about the height trajectory of the UFO.

Solutions. No, because this GP ascribes a non-zero probability to the event that the height is negative.

- (ii) This time, consider our a priori knowledge about the log-height $\ell(t) = \log(h(t))$ is modelled by supposing a priori that $\ell \sim \mathcal{GP}(\mu, k)$. Subsequently, we receive the observation $h(\xi) = \exp(6)$ for time $\xi = 2$ which we use to create our data set D .

Give the formulae for the posterior mean function $m_{\ell|D}(t) = \mathbb{E}[L_t|D]$ and variance function $v_{\ell|D}(t) = \mathbb{V}[L_t|D]$ for the log-height random variable $L_t = \ell(t)$ at time t and evaluate them at time $t = 3$ (give an approximate numerical value for each).

Solutions. First, what is asked for is the posterior mean function $m_{\ell|D}(t)$ for $t = 2$ and $D = \{(0, 1)\}$ We utilise the formula for the posterior mean function $m_{\ell|D}$ of a GP to get

$$\begin{aligned} m_{\ell|D}(t) &= \mu(t) + k(t, \xi)k(\xi, \xi)^{-1}(\ell(\xi) - \mu(\xi)) \\ &= 10 - 4k(t, \xi) = 10 - 4 \exp(-4|t - 2|^2). \end{aligned}$$

Hence, $m_{\ell|D}(3) = 10 - 4e^{-4} \approx 10 - 4 * 0.0183 \approx 9.926$.

Similarly, we have

$$\begin{aligned} v_{\ell|D}(t) &= k(t, t) - k(t, \xi)k(\xi, \xi)^{-1}k(\xi, t) \\ &= 1 - \exp(-8|t - 2|^2); \end{aligned}$$

Hence,

$$v_{\ell|D}(3) = 1 - e^{-8} \approx 0.9996.$$

- (iii) Let H_t denote the random variable of the height $h(t)$ at time t after having observed the data D . Conditional on the data, work out a 95 % confidence interval of the height in the infinite time limit. That is, specify $a, b \in \mathbb{R}$ such that $Pr[\lim_{t \rightarrow \infty} H_t \in [a, b]] = 95\%$.

HINT: You may make use of the following facts:

- For any Gaussian distributed random variable $X \sim \mathcal{N}(m, \sigma^2)$ a 95% confidence interval is given by $[m - 1.96\sigma, m + 1.96\sigma]$.
- Suppose $(X_t)_{t \in \mathbb{R}}$ is a sequence of Gaussian random variables with $X_t \sim \mathcal{N}(m_t, \sigma_t^2)$ for all t and where the moments converge: $\lim_{t \rightarrow \infty} m_t =: m_\infty \in \mathbb{R}$, $\lim_{t \rightarrow \infty} \sigma_t =: \sigma_\infty \in \mathbb{R}$. Then we can define the limit $\lim_{t \rightarrow \infty} X_t =: X_\infty$ as a random variable with distribution $\mathcal{N}(m_\infty, \sigma_\infty^2)$.

Solutions. We notice that if $[u, v]$ is a confidence interval of the normally distributed r.v. L_t then, because $\exp(\cdot)$ is a strictly monotonically increasing function (and an injection), that $[a, b]$ with $a = \exp(u)$ and $b = \exp(v)$ is a confidence interval with the same level of confidence for the r.v. $H_t = \exp(L_t)$. Hence, all we need to work out are u, v for the r.v.

$$L_\infty = \lim_{t \rightarrow \infty} L_t \sim \mathcal{N}(m_\infty, v_\infty)$$

where

$$m_\infty = \lim_{t \rightarrow \infty} m_{\ell|D}(t) = 10$$

and

$$v_\infty = \lim_{t \rightarrow \infty} v_{\ell|D}(t) = 1.$$

By the hint, we can define

$$u := 10 - 1.96, \quad v = 10 + 1.96$$

Hence $a = \exp(u) \approx 3102.61$ and $b = \exp(v) \approx 156373.084$

[9 marks]

- (b) In the lectures you have been told that the sum of two covariance functions is a covariance function. Now we ask you to show it in the following scenario: Suppose you have two valid continuous covariance functions $k_1, k_2 : \mathbb{X} \rightarrow \mathbb{R}$ on the domain $\mathbb{X} = [0, 1]^2 \subset \mathbb{R}^2$. Show that on \mathbb{X} the function $k = k_1 + k_2$ is also a valid covariance function.

[4 marks]

Solutions.

We appeal to Mercer's condition showing that $\langle f, T_k f \rangle > 0 \forall f \in L_2([0, 1]^2)$. Since k_1, k_2 are continuous covariance functions on compact support, Mercer's condition holds and hence $\langle f, T_{k_1} f \rangle, \langle f, T_{k_2} f \rangle > 0 \forall f \in L_2([0, 1]^2)$. Now for all f :

$$\begin{aligned} \langle f, T_k f \rangle &= \int_0^1 \int_0^1 f(x)(k_1(x, x') + k_2(x, x'))f(x') dx dx' \\ &= \int_0^1 \int_0^1 (f(x)k_1(x, x')f(x') + f(x)k_2(x, x')f(x')) dx dx' \\ &= \underbrace{\int_0^1 \int_0^1 f(x)k_1(x, x')f(x') dx dx'}_{\langle f, T_{k_1} f \rangle > 0} + \underbrace{\int_0^1 \int_0^1 f(x)k_2(x, x')f(x') dx dx'}_{\langle f, T_{k_2} f \rangle > 0} > 0 \end{aligned}$$

q.e.d.

- (c) Consider a distribution $P(A, B, C, D, E)$ of the discrete random variables A, B, C, D, E and suppose the distribution has a conditional independence structure admitting the factorisation $P(A, B, C, D) = P(A|B, C)P(C)P(D|C)P(B)P(E)$. Draw a Bayes net represent-

ing this conditional independence structure.

[3 marks]

