PATTERN FORMATION

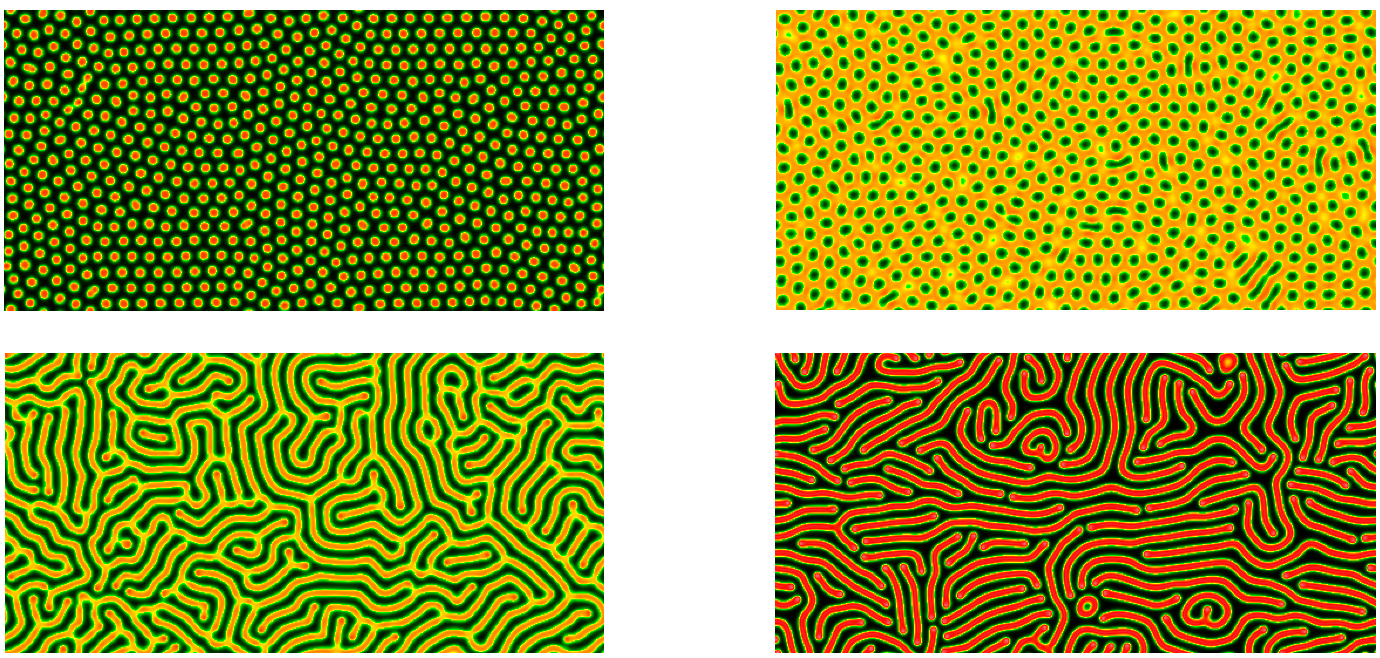
Il fenomeno dei pattern è molto frequente in natura, soprattutto nei processi di morfogenesi. Un primo studio dei pattern fu effettuato da A. Turing nel 1952. L'idea di base è quella di modellizzare l'interazione chimica tra 2 popolazioni di molecole, per esempio:

$$\boxed{\begin{cases} \partial_t u = f(u,v) + D_u \Delta u \\ \partial_t v = g(u,v) + D_v \Delta v \end{cases}}$$

L'apparizione spontanea di questi fenomeni è caratterizzato dalle BIFORCAZIONI DI TURING. Un esempio di modello di questo tipo è il MODELLO DI GRAY-SCOTT:

$$\boxed{\begin{cases} \partial_t u = F\cdot(1-u) - uv^2 + D_u \Delta u \\ \partial_t v = -(F+K)v + uv^2 + D_v \Delta v \end{cases}}$$

con $F, K \in \mathbb{R}$



IDENTIFICAZIONE DEI PARAMETRI

È possibile, a partire da alcune osservazioni di un fenomeno fisico/biologico, identificare i parametri di un modello associato a tale fenomeno?

Consideriamo un modello ODE:

$$\boxed{\dot{x} = F(x, \Theta)}$$

con $x \in \mathbb{R}^n$, $\Theta \in \Omega \subseteq \mathbb{R}^p$, $F: \mathbb{R}^n \times \Omega \to \mathbb{R}^n$

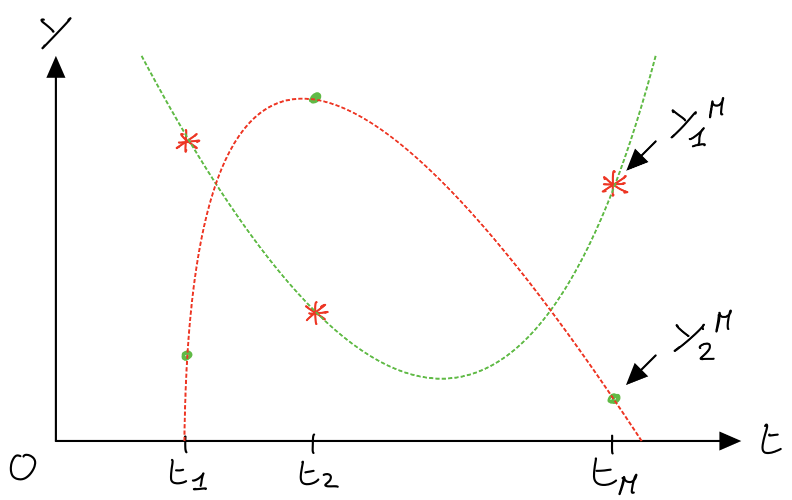
$\Rightarrow$ vogliamo caratterizzare il vettore dei parametri Θ a partire dalle osservazioni sperimentali della s-sima quantità al tempo m Y_s^m:

$$Y_s^m = H_s(x(t_m), \Theta) + \varepsilon_s^m \qquad \begin{matrix} s = 1, \dots, V \\ m = 1, \dots, M \end{matrix}$$

dove:

$H: \mathbb{R}^n \times \Omega \to \mathbb{R}^n$ è la funzione di stato del sistema

$\varepsilon_s^m \in \mathbb{R}^n$ è il vettore degli errori della misurazione



$\Rightarrow$ vogliamo trovare le traiettorie che più si avvicinano all'andamento delle Y_s^m in funzione dei parametri Θ, ovvero a partire da $\{y_s^m\}_{\substack{s=1,\dots,v \\ m=1,\dots,M}}$ vogliamo ricostruire il vettore $\Theta \in \Omega$ utilizzando tecniche del tipo MINIMI QUADRATI

Minimi Quadrati:

Definiamo $L(\Theta) := \sum_{s=1}^{v} \sum_{m=1}^{M} [y_s^m - H_s(x(t_m), \Theta)]^2 = \sum_{s=1}^{v} \sum_{m=1}^{M} (\mathcal{E}_s^m)^2$

N.B.

L'applicazione H é, in generale, NON lineare !!!

esempio (Modello su infezione di HIV):

Siano:

$z(t) :=$ popolazione di infetti da HIV asintomatici,

$x(t) :=$ popolazione di infetti da HIV sintomatici

$\Rightarrow$ vogliamo stimare la velocitá di propagazione dei sintomi da HIV $v(t)$:

$$\boxed{z(t)} \xrightarrow{v(t)} \boxed{x(t)}$$

Il modello che consideriamo é:

$$\boxed{\begin{cases} \dot{z} = -v(t)z, & z(0) = 1 \\ \dot{x} = v(t)z, & x(0) = 0 \end{cases}}$$

$\Rightarrow$ assumiamo $v(t) = at, \quad a > 0$

$\Rightarrow \dot{z} = -atz \Rightarrow z(t) = z(0)e^{-a\frac{t^2}{2}}, \quad x(t) = 1 - z(t)$

Assumiamo di conoscere $\left\{\frac{d}{dt}x(t_m)\right\}_{m=1}^{M} \approx \{y^m\}_{m=1}^{M}$. Si ha quindi:

$$L(a) = \sum_{m=1}^{M} \left[\frac{d}{dt}x(t_m) - y_m\right]$$

Stimiamo a calcolando $\hat{a} = \operatorname{argmin} L(a)$ ovvero:

$$L(a) = \sum_{m=1}^{M} \left[a\, t_m e^{-a\frac{t_m^2}{2}} - y_m \right]^2, \quad a \in \mathbb{R}^{\geq 0}$$

$$\Rightarrow \quad 0 = L'(a) = \sum_{m=1}^{M} 2\left(a - \frac{t_m^2}{2}\right) e^{-a\frac{t_m^2}{m}} \left(a\, t_m e^{-a\frac{t_m^2}{2}} - y_m\right)$$

(In MATLAB si utilizza bisezione etc...)

$\Rightarrow$ si tratta di una minimizzazione vincolata.

Minimi Quadrati Generalizzati:

Date $\{y_s^m\}_{\substack{s=1,\dots,v \\ m=1,\dots,M}}$ le misurazioni con errori $\varepsilon_s^m \in \mathbb{R}^v$, assumiamo di avere:

$$\mathbb{E}\left(\varepsilon^m \cdot (\varepsilon^n)^T\right) = \begin{cases} 0 & \text{se } m \neq n \\ \Sigma_m \in \mathbb{R}^{v \times v} & \text{se } m = n \end{cases}$$

dove:

$$\Sigma_m := \begin{pmatrix} \sigma_1^2 & \cdots & Cov(\varepsilon_i, \varepsilon_s) \\ Cov(\varepsilon_i, \varepsilon_s) & \cdots & \sigma_v^2 \end{pmatrix}$$ é la MATRICE DI COVARIANZA

esempio:

Se gli errori $\varepsilon_i^m, \varepsilon_s^m$ sono indipendenti $\forall i, s$, allora Σ_m é diagonale.

$\Rightarrow$ Come influisce l'incertezza in y sulla stima di $\hat{\theta}$?

Definiamo l'errore dei MQG:

$$L(\theta) = \sum_{m=1}^{M} \| y^m - H(x(t_m), \theta) \|_{W_m}^2$$

dove:

$$\|v\|_{W_m}^2 = v^T W_m v, \quad v \in \mathbb{R}^v, \quad W_m \in \mathbb{R}^{v \times v} \text{ def. pos.}$$

Consideriamo θ^* (la STIMA di θ) e associamo un

INTERVALLO DI CONFIDENZA $I_\alpha(\theta^*)$ che misura l'affidabilità della stima di θ con $\alpha :=$ misura dell'accuratezza della stima θ^*

<u>esempio:</u>

$x = (x_1, \dots, x_n)$ campione di n elementi, realizzazioni di una variabile aleatoria Gaussiana $X \sim N(\mu, \sigma^2)$ di cui μ NON é nota, mentre σ^2 é nota.

$\Rightarrow$ vogliamo determinare una stima per la media $\hat{\mu}$ e anche l'intervallo di confidenza $I_\alpha(\hat{\mu})$.

Assumiamo che $\mathbb{P}(\mu \in I_\alpha(\hat{\mu})) = 1 - \alpha$

$\Rightarrow$ con "confidenza" $1-\alpha$ il parametro μ sará incluso in $I_\alpha(\hat{\mu}) = [\hat{\mu} - c_\alpha, \hat{\mu} + c_\alpha]$

$\Rightarrow$ si ha:

$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i \Rightarrow \mathbb{E}(\hat{\mu}) = \mu$ ($\hat{\mu}$ é uno **STIMATORE NON DISTORTO**)

(per la varianza si ha:

$$\hat{\sigma}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\mu})^2 \approx \mathbb{E}[(X - \hat{\mu})^2] =: Var(X))$$

Quindi $\hat{\mu} \sim N(\mu, \frac{\sigma^2}{n})$ e riscalando otteniamo:

$$\frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \Rightarrow \mathbb{P}\left(-z_{(1-\frac{\alpha}{2})} \leq \frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}} \leq z_{(1-\frac{\alpha}{2})}\right) = 1 - \alpha$$

$\left(z_{(1-\frac{\alpha}{2})} = \phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right)$

Quindi $c_\alpha = z_{(1-\frac{\alpha}{2})} \cdot \frac{\sigma}{\sqrt{n}}$

<u>Analisi della Sensitività:</u>

Dato il modello $\dot{x} = F(x, \theta)$, $x \in \mathbb{R}^n$, $\theta \in \mathbb{R}^p$, introduciamo la MATRICE DI SENSITIVITÁ:

$$D(t) = \begin{pmatrix} \partial_{\theta_1} x_1 & \cdots & \partial_{\theta_p} x_1 \\ \vdots & & \vdots \\ \partial_{\theta_1} x_n & \cdots & \partial_{\theta_p} x_n \end{pmatrix} \in \mathbb{R}^{n \times p}$$

$\Rightarrow$ $D(t)$ determina come i parametri impattano sulle singole variabili:

$$\frac{d}{dt}\left(\partial_{\theta_s} x_i\right) = \sum_{k=1}^{n} \partial_{x_k} F_i(x_i, \theta)\, \partial_{\theta_s} x_k + \partial_{\theta_s} F_i(x_i, \theta), \quad \begin{matrix} i = 1, \dots, n \\ s = 1, \dots, p \end{matrix}$$

ovvero:

$$A = (A_{ik}) = \left(\partial_{x_k} F_i(x_i, \theta)\right), \quad B = (B_{is}) = \left(\partial_{\theta_s} F_i(x_i, \theta)\right)$$

$\Rightarrow$ in forma matriciale si ha il seguente sistema:

$$\boxed{\frac{d}{dt} D(t) = A(t) \cdot D(t) + B(t)}$$

($n \times p$ equazioni differenziali)

con dato iniziale $D(0) = 0$

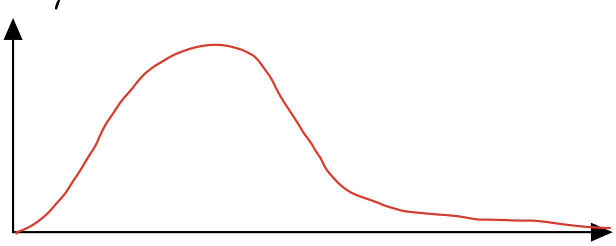
<u>esempio</u>:

$\dot{x} = -\theta x, \quad x(0) = 1$

la sol. é data da $x(t) = x(0) e^{-\theta t}$, quindi:

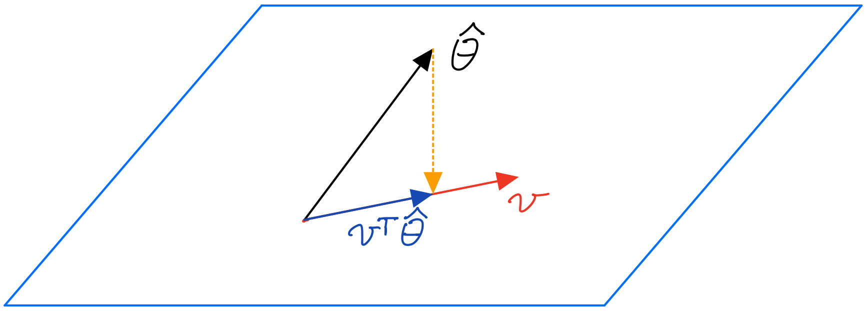
$$\partial_\theta x = -t e^{-\theta t}$$

$\Rightarrow$ graficando la funzione $|\partial_\theta x|$ si ha:



Intervallo di Confidenza Multivariato:

Dato $\hat{\theta} \in \mathbb{R}^p$ stimatore del vettore $\theta \in \mathbb{R}^p$ vogliamo definire $I_\alpha(\hat{\theta})[v]$, ovvero l'intervallo di confidenza applicato al vettore v con $\|v\| = 1$:



Se $\hat{\theta} \sim f_{\hat{\theta}}(s)$ (distribuzione di una variabile aleatoria multivariata) si ha:

$$I_\alpha(\theta)[v] = [v^T\hat{\theta} - C_\alpha[v],\ v^T\hat{\theta} + C_\alpha[v]]$$
$$\Rightarrow \mathbb{P}(v^T\hat{\theta} \in I_\alpha(\theta)[v]) = 1 - \alpha$$

esempio:

Se $\hat{\theta} \sim N(0, \Sigma)$ con $f_{\hat{\theta}}(s) = \frac{1}{\sqrt{(2\pi)^p \cdot \det(\Sigma)}} e^{\left(-\frac{1}{2} s^T \Sigma^{-1} s\right)}$

(ovvero $\hat{\theta}$ é NORMALE MULTIVARIATA):

$\Rightarrow$ introduciamo $e_i = (0, \dots, 0, \overset{i\text{-sima}}{1}, 0, \dots, 0)$:

$e_i^T\hat{\theta} = \hat{\theta}_i \Rightarrow I_\alpha(e_i^T\hat{\theta})$ é l'intervallo di confidenza associato alla componente i-sima.

$$\mathbb{P}(e_i^T\hat{\theta} \in I_\alpha(e_i\hat{\theta})) = \mathbb{P}\left(-z_{(1-\frac{\alpha}{2})} \leq \frac{e_i^T(\hat{\theta} - \theta)}{\sqrt{e_i^T \Sigma e_i}} \leq z_{(1-\frac{\alpha}{2})}\right)$$

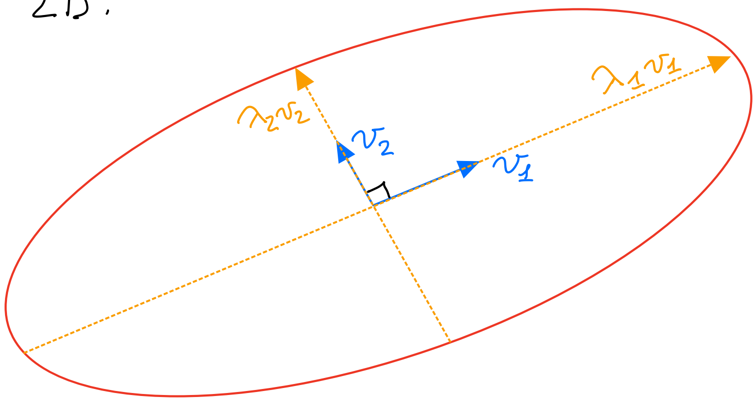
$\Rightarrow$ l'intervallo di confidenza é:

$$I_\alpha(\hat{\theta})[e_i] = [e_i^T\hat{\theta} - z_{(1-\frac{\alpha}{2})}\sigma_i,\ e_i^T\hat{\theta} + z_{(1-\frac{\alpha}{2})}\sigma_i]$$

Assumiamo che Σ sia def. pos., allora i suoi autoval. sono $\lambda_i(\Sigma) > 0 \ \forall i = 1, \dots, p$ e i suoi autovettori v_i

sono ortogonali. Definiamo $\mathcal{E} := z^T \Sigma^{-1} z$, $z \in \mathbb{R}^p$ l'ellissoide t.c. v_i sono le direzioni degli assi e λ_i sono le lunghezze dei diametri.

caso 2D:



Quantificazione dell'incertezza:

Ricapitolando, abbiamo:

$$\hat{\theta} = \operatorname{argmin} L(\theta) = \sum_{m=1}^{M} \| y^m - H(x(t_m), \theta) \|^2_{W_m}$$

Assumiamo che $\varepsilon^m \sim N(0, \Sigma_m)$, qual è l'impatto che ha ε^m su $\hat{\theta}$?

$\Rightarrow$ linearizziamo: $L(\theta + \xi)$, $\xi = \hat{\theta} - \theta \in \mathbb{R}^p$

$$\Rightarrow H_i(x^m, \theta + \xi) \simeq H_i(x^m) + \left[\sum_{s=1}^{n} \sum_{\ell=1}^{p} \underbrace{\partial_{x_s} H_i(x^m)}_{R_{is}(t_m)} \underbrace{\partial_{\theta_\ell} x_s^m}_{D_{s\ell}(t_m)} \right] \cdot \xi$$

In forma vettoriale:

$$H(x^m, \theta + \xi) \simeq H(x^m) + R_m \cdot D_m \cdot \xi$$

dove:

$$R_m = \left(R_{is}(t_m) \right)_{\substack{i=1,\dots,r \\ s=1,\dots,n}}, \quad D_m := \text{matrice di sensitività}$$

$$\Rightarrow L(\theta + \xi) = \sum_{m=1}^{M} \| y^m - H(x^m, \theta + \xi) \|^2_{W_m}$$

$$\Rightarrow y^m = H(x^m) + \varepsilon^m + \text{linearizzazione } H$$

$$= \sum_{m=1}^{M} \| \varepsilon^m - R_m D_m \xi \|^2_{W_m}$$

$\Rightarrow \hat{\theta}$ (linearizzato) sarà soluzione di:

$$0 = \nabla_\xi L(\underbrace{\theta + \xi}_{\hat{\theta}}) = \sum_{\mu=1}^{M} \left(-D_\mu^T R_\mu^T W_\mu \cdot \varepsilon^\mu\right) + \sum_{\mu=1}^{M} \left(D_\mu^T R_\mu^T W_\mu R_\mu D_\mu \xi\right)$$

$\Rightarrow$ definiamo $b(\varepsilon) = \sum_{\mu=1}^{M} \left(-D_\mu^T R_\mu^T W_\mu \cdot \varepsilon^\mu\right)$,

$G \cdot \xi = \sum_{\mu=1}^{M} \left(D_\mu^T R_\mu^T W_\mu R_\mu D_\mu \xi\right)$, si ha il seguente sistema:

$$G \cdot \xi = b(\varepsilon)$$

$\Rightarrow$ se $\det G \neq 0$, allora $\xi = G^{-1} b(\varepsilon)$ e tale quantità quantifica l'errore nella stima $\hat{\theta}$.

Quindi possiamo calcolare la seguente quantità:

$V = \mathbb{E}(\xi \cdot \xi^T)$ la MATRICE DI COVARIANZA DI ξ

usando la precedente si ha:

$$\mathbb{E}(G \cdot \xi \cdot \xi^T G^T) = G\,\mathbb{E}(\xi \cdot \xi^T) = G V G^T$$

$$\mathbb{E}\left(b(\varepsilon)\, b(\varepsilon)^T\right) = \mathbb{E}\left(\sum_{\mu=1} D_\mu^T R_\mu^T W_\mu \cdot \underbrace{\varepsilon^\mu (\varepsilon^\mu)^T}_{\Sigma_\mu} W_\mu^T R_\mu D_\mu\right)$$

quindi V è identificata univocamente dalle matrici di covarianza $\{\Sigma_\mu\}_{\mu=1}^{M}$ se G è NON singolare.

$\Rightarrow$ dato che $\varepsilon^\mu \sim N(0, \Sigma_\mu)$, anche $\xi \sim N(0, \Sigma_\mu)$

Osservazione:

L'equazione $\varepsilon = z^T V^{-1} z$ definisce l'iper-ellissoide di incertezza nella stima $\hat{\theta}$:

$$\sqrt{\lambda_{MIN}(V)} \leq \tau_i \leq \sqrt{\lambda_{MAX}(V)} \qquad i = 1, \dots, p$$

Condizionamento della stima:

$$\frac{\lambda_{MAX}}{\lambda_{MIN}} = \operatorname{cond}(P)$$

$\Rightarrow$ maggiore è cond(P), maggiore sarà un autoval. rispetto agli altri.

$\Rightarrow$ l'intervallo di confidenza di alcuni parametri sarà meno affidabile.

<u>Scelta di W_m:</u>

$b(\varepsilon)\, b(\varepsilon)^T = \sum_{m=1}^{M} D_m^T R_m^T W_m \Sigma_m W_m^T R_m D_m$,

$G V G^T$ dove $G = \sum_{m=1}^{M} D_m^T R_m^T W_m$

$\Rightarrow$ la scelta ottimale per W_m è $W_m = \Sigma_m^{-1}$,

con tale scelta si ha $V = G^{-1}$ (se $\det G \neq 0$)

<u>esempio:</u>

$\dot{x} = -\theta x$, $x(0) = 1$, $t \in [0, 1]$.

$L(\theta) = \frac{1}{\sigma^2} |\underbrace{y(1)}_{\text{misurazione in } t=1} - x(1)|^2$, $m = M$

Quindi:

$H(x) = x$; $W_M = \frac{1}{\sigma^2}$

$\Rightarrow$ l'equazione di sensitività è $x(t) = e^{-\theta t}$

$\Rightarrow D(t) = -t e^{-\theta t}$, $D(1) = -e^{-\theta}$

$\Rightarrow R(1) = 1$ $(\equiv \partial_x x(1))$

Quindi:

$$G(1) = \partial_\theta x \cdot \partial_x x \cdot W \cdot \partial_x x\, \partial_\theta x \Big|_{t=1} = \frac{1}{\sigma^2} e^{-2\theta}$$

$$\mathbb{E}(b(\varepsilon)\, b(\varepsilon)^T) = \partial_\theta x \cdot \partial_x x \cdot W \cdot \Sigma \cdot W \cdot \partial_x x \cdot \partial_\theta x$$

$$= t^2 \frac{1}{\sigma^2} e^{-2\theta t}$$

$$\mathbb{E}(G V G^T) = t^2 \frac{1}{\sigma^2} e^{-2\theta t}$$

$$V = G^{-1} = e^{2\theta}\frac{\tau^2}{t^2} \equiv \tau_\theta^2$$