Stochastic modeling and estimation of tumor growth under immunotherapy

Modibo DIABATE (LJK-UGA), Loren COQUILLE (IF-UGA), Adeline SAMSON (LJK-UGA)

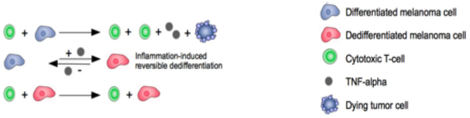
Ecole de printemps de la Chaire MMB

Aussois, May 2019

Outline

Immunotherapy

- Context
  - stochastic model for skin cancer **immunotherapy** (Baar et al., 2015)
  - Adoptive Cell Transfer (ACT) therapy (Landsberg et al., 2012)

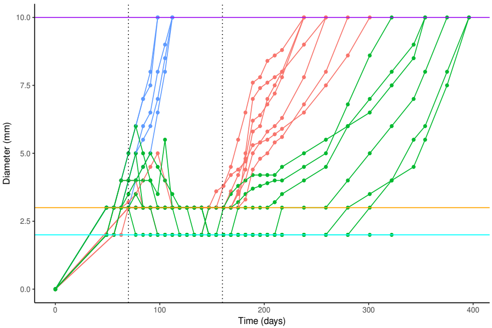


Objective: understanding the resistance of tumors & treatment optimization

- Parameter estimation using real data
- Evaluation of T cell exhaustion probability
- Treatment optimization to delay relapse

Experimental data

- Three groups of mice
  - CTRL mice (5): no therapy
  - ACT mice (7): ACT therapy (70th day)
  - ACT+Re mice (7): ACT therapy (70th, 160th day)



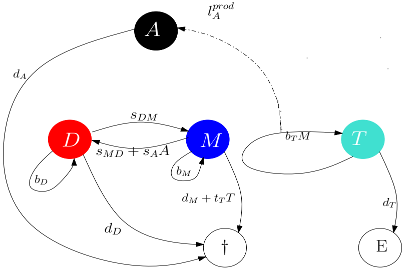
- Tumor evolution measurement

$$y_{ij}^{obs} = \begin{cases} y_{ij} \text{ if } y_{ij} \in]3 \text{ mm}, 10 \text{ mm}[\\ 1.9 \text{ mm if } y_{ij} \in [0 \text{ mm}, 1.9 \text{ mm} \\ 3 \text{ mm if } y_{ij} \in [2 \text{ mm}, 3 \text{ mm}] \\ 10 \text{ mm if } y_{ij} \geq 10 \text{ mm} \end{cases}$$

Stochastic model: continuous time Markov process (Baar et al., 2015)

- $\bullet$: $M(t)$: population of differentiated cells at time t;
- $\bullet$: $D(t)$: population of dedifferentiated cells;
- $\bullet$: $T(t)$: population of therapy cells;
- $\bullet$: $A(t)$: population of cytokines TNF_α.

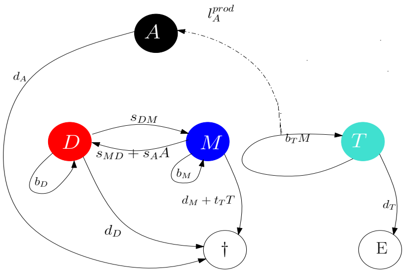
  $X(t) = (M(t), D(t), T(t), A(t)) \in \mathbb{N}^4$



Stochastic model: continuous time Markov process (Baar et al., 2015)

- $\bullet$: $M(t)$: population of differentiated cells at time t;
- $\bullet$: $D(t)$: population of dedifferentiated cells;
- $\bullet$: $T(t)$: population of therapy cells;
- $\bullet$: $A(t)$: population of cytokines TNF_α.

$$X(t) = (M(t), D(t), T(t), A(t)) \in \mathbb{N}^4$$



Stochastic model = Birth and Death Process + switch + Predator-Prey

Birth and Death Process

$$\begin{cases} \bullet \rightarrow \bullet\bullet & \text{with a rate } b \\ \bullet \rightarrow \dagger & \text{with a rate } d \end{cases}$$

Predator-Prey system

$$\begin{cases} \dot{M}(t) = M(t)\,(\alpha - \beta T(t)) \\ \dot{T}(t) = T(t)\,(\delta M(t) - \gamma) \end{cases}$$

$$\theta' = (b_M, d_M, b_D, d_D, s_{MD}, s_{DM}, b_T, d_T, t_T, s_A, d_A, l_A^{\text{prod}})$$

Deterministic approximation of cancer model

Large population approximation: **Stochastic model $\rightarrow$ Deterministic model**

Let $M_K(t) = \frac{M(t)}{K}$, $D_K(t) = \frac{D(t)}{K}$, $T_K(t) = \frac{T(t)}{K}$, $A_K(t) = \frac{A(t)}{K}$

$$\lim_{K\to\infty} (M_K(t), D_K(t), T_K(t), A_K(t)) \stackrel{d}{=} (n_M^*, n_D^*, n_T^*, n_A^*)$$

where $(n_M^*, n_D^*, n_T^*, n_A^*)$ is the solution of the deterministic system:

$$\begin{cases} \dot{\mathbf{n}}_\mathbf{M} = \overbrace{(b_M - d_M)}^{r_M}\mathbf{n}_\mathbf{M} - t_T\mathbf{n}_\mathbf{T}\mathbf{n}_\mathbf{M} - s_{MD}\mathbf{n}_\mathbf{M} + s_{DM}\mathbf{n}_\mathbf{D} - s_A\mathbf{n}_\mathbf{A}\mathbf{n}_\mathbf{M} \\ \dot{\mathbf{n}}_\mathbf{D} = \underbrace{(b_D - d_D)}_{r_D}\mathbf{n}_\mathbf{D} + s_{MD}\mathbf{n}_\mathbf{M} - s_{DM}\mathbf{n}_\mathbf{D} + s_A\mathbf{n}_\mathbf{A}\mathbf{n}_\mathbf{M} \\ \dot{\mathbf{n}}_\mathbf{T} = -d_T\mathbf{n}_\mathbf{T} + b_T\mathbf{n}_\mathbf{M}\mathbf{n}_\mathbf{T} \\ \dot{\mathbf{n}}_\mathbf{A} = -d_A\mathbf{n}_\mathbf{A} + l_A^{prod} b_T\mathbf{n}_\mathbf{M}\mathbf{n}_\mathbf{T} \end{cases}$$

with initial condition $(n_{M_0}, n_{D_0}, n_{T_0}, n_{A_0})$

$\theta' = (r_M, r_D, s_{MD}, s_{DM}, n_{M_0}, b_T, d_T, t_T, s_A, d_A, l_A^{\text{prod}})$ (to be estimated)

Parameter estimation in deterministic approximation: Nonlinear Mixed Effects Model (Pinheiro and Bates, 1995)

$$y_{ij} = f(\zeta_i, t_{ij}) + \epsilon_{ij},\ \epsilon_i \sim \mathcal{N}(0, \sigma^2 I_{n_i}),$$

$$\log(\zeta_i) = \log(\theta' \odot d_i)\ +\ r_i \odot d_i,\ r_i \sim \mathcal{N}(0, \Omega)$$

$f(\zeta_i, t) = (n_M^*(\zeta_i, t) + n_D^*(\zeta_i, t))^{\frac{1}{3}}$: tumor size (n_M^*, n_D^* solution of the deterministic system);

$\epsilon_i = (\epsilon_{i1}, \ldots, \epsilon_{in_i})$: residual error;

ζ_i : individual parameters;

$\theta' = (r_M, r_D, s_{MD}, s_{DM}, n_{M_0}, b_T, d_T, t_T, s_A, d_A, l_A^{\text{prod}})$: vector of fixed effects;

$d_i = (1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0)$: when i is a control individual;

$d_i = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$: when i is a treated individual;

$r_i = (r^i_{r_M}, r^i_{r_D}, r^i_{s_{MD}}, r^i_{s_{DM}}, r^i_{n_{M_0}}, r^i_{b_T}, r^i_{d_T}, r^i_{t_T}, r^i_{s_A}, r^i_{d_A}, r^i_{l_A})$: random effects,

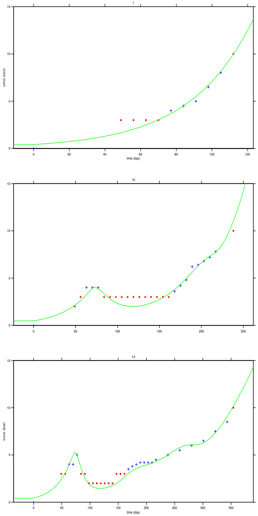
Ω : variance matrix of the random effects;

$\varrho = \{\theta', \sigma, \Omega\} \in \Theta$

Estimation

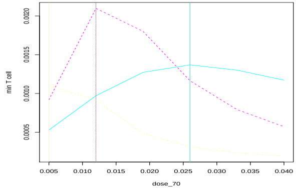
- SAEM (Stochastic Approximation Expectation Maximization) for censored data
- Model selection through likelihood ratio tests

	mean (r.s.e (%))	sd (r.s.e (%))
r_M	0.09 (4)	0.10 (34)
r_D	0.05 (10)	0.35 (20)
s_{MD}	< 0.01 (118)	0 (-)
s_{DM_p}	< 0.01 (-)	0 (-)
n_{M_0}	0.08 (28)	0.58 (29)
b_T	< 0.01 (16)	0 (-)
d_T	0.02 (34)	1.07 (27)
t_T	1.33 (55)	1.03 (48)
s_A	77 (37)	0 (-)
d_A	0.03 (-)	0.09 (181)
l_A^{prod}	0.19 (93)	3.23 (25)
σ	0.44 (6)	

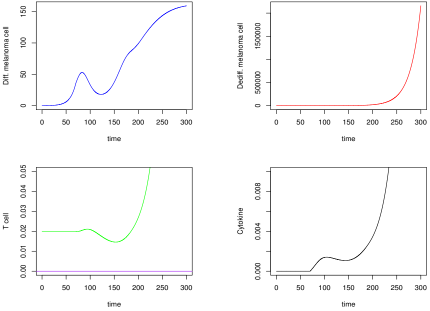


Using estimated parameters, we:

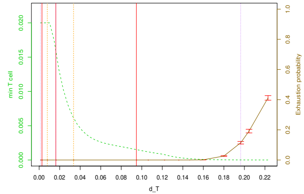
- Simulate T cell exhaustion (Relapse)
- Estimate T cell exhaustion probability
- Optimize treatment doses and times



Treatment optimization



Cell populations dynamics



Exhaustion proba. vs d_T (Imp. Splitting)

Stochastic diffusion approximation of cancer Model (Cseke et a.l, 2016)

Cancer model: continuous time discrete state Markov Jump Process

$$S = \begin{bmatrix} 1 & -1 & 0 & -1 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & l_A^{\text{prod}} & 0 & -1 \end{bmatrix}, \quad g(X(t), \theta') = \begin{bmatrix} r_M M(t) \\ t_T M(t) T(t) \\ r_D D(t) \\ s_{MD} M(t) \\ s_{DM} D(t) \\ s_A M(t) A(t) \\ b_T M(t) T(t) \\ d_T T(t) \\ d_A A(t) \end{bmatrix}$$

$X(t) = (M(t), D(t), T(t), A(t)) \in \mathbb{N}^4$: state of the stochastic model at time t

$X(t) \to X(t) + S_{\cdot r}$: one possible transition with rate $g_r(X(t), \theta')$, $r \in \{1, \dots, R = 9\}$

$x_t = (x_{M_t}, x_{D_t}, x_{T_t}, x_{A_t}) \in \mathbb{R}^4$: continuous real-valued pendant to $\frac{X(t)}{K}$

Cancer diffusion process:

$$a(x_t, \theta') = S g(x_t, \theta') \quad \text{and} \quad b(x_t, \theta') = S \text{diag}(g(x_t, \theta')) S^T$$

$$dx_t = a(x_t, \theta') dt + b(x_t, \theta')^{\frac{1}{2}} dW_t.$$

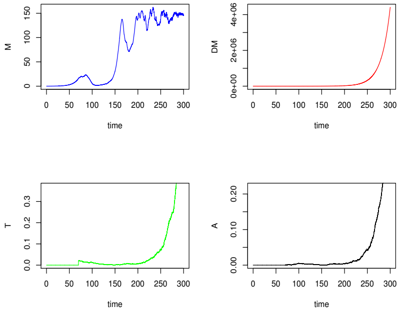
Stochastic diffusion approximation of cancer model

Diffusion process:

$$dx_t = a(x_t, \theta')dt + b(x_t, \theta')^{\frac{1}{2}} dW_t, \quad x_{t_0} = x_{\text{initial}}$$

Euler discretization: $x_{t_i} = x_{t_{i-1}} + a(x_{t_{i-1}}, \theta')\Delta_i + b(x_{t_{i-1}}, \theta')^{\frac{1}{2}} \rho \sqrt{\Delta_i},$

$$\rho = (\rho_1, \rho_2, \rho_3, \rho_4), \ \rho_l \sim \mathcal{N}(0,1), \ l = 1, \dots, 4$$



Simulation using the previous estimated value of θ'

Parameter estimation for diffusion process (In progress . . .)

Diffusion process:

$$dx_t = a(x_t, \theta')dt + b(x_t, \theta')dW_t,$$

Noisy observations $y = (y_1, \ldots, y_n)$ of x at time t_i, $i = 1, \ldots, n$:

$$y_i = H(x_i) + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \varsigma^2),\ \theta = \{\theta', \varsigma\}.$$

Strategy: Calculate the likelihood $p(y|\theta) = \int p(x, y|\theta)dx$ and compute

$$\theta^{\text{opt}} = \arg\max_{\theta} p(y|\theta)$$

Parameter estimation for diffusion process (In progress ...)

Diffusion process:

$$dx_t = a(x_t, \theta')dt + b(x_t, \theta')dW_t,$$

Noisy observations $y = (y_1, \ldots, y_n)$ of x at time t_i, $i = 1, \ldots, n$:

$$y_i = H(x_i) + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \varsigma^2), \ \theta = \{\theta', \varsigma\}.$$

Strategy: Calculate the likelihood $p(y|\theta) = \displaystyle\int p(x, y|\theta)dx$ and compute

$$\theta^{\text{opt}} = \arg\max_{\theta} p(y|\theta)$$

Likelihood calculation $\rightarrow$ approximate inference methods ($p(y|\theta) \approx \tilde{p}(y|\theta)$)

- Expectation Propagation (EP) method (Minka, 2001; Heskes and Zoeter, 2002)
  - ◇ Approximate p (factorizable) by q (factorizable) by minimizing $DKL(p||q)$
  - ◇ Good compromise between accuracy and speed (Barthelmé and Chopin, 2014)

Conclusion

- Done
  - Simplification/adaptation of the cancer model for parameter estimation
  - Parameter estimation with NLMEMs and SAEM algorithm using (censored) real data
  - T cell exhaustion probability estimation
  - Treatment optimization
  - Parameter estimation using EP in AR(1) and Ornstein Uhlenbeck latent models
- In progress & Future Works
  - Parameter estimation in cancer diffusion approximation using EP method
  - Extend the calculations to diffusion model with random effects

References

Martina Baar, Loren Coquille, Hannah Mayer, Michael Hölzel, Meri Rogava, Thomas Tüting, and Anton Bovier. A stochastic individual-based model for immunotherapy of cancer. *Scientific Reports*, 2015.

Simon Barthelmé and Nicolas Chopin. Expectation propagation for likelihood-free inference. *Journal of the American Statistical Association*, 109(505):315–333, 2014.

Tom Heskes and Onno Zoeter. Expectation propagation for approximate inference in dynamic bayesian networks. In *Proceedings of the Eighteenth conference on Uncertainty in artificial intelligence*, pages 216–223. Morgan Kaufmann Publishers Inc., 2002.

Jennifer Landsberg, Judith Kohlmeyer, Marcel Renn, Tobias Bald, Meri Rogava, Mira Cron, Martina Fatho, Volker Lennerz, Thomas Wölfel, Michael Hölzel, et al. Melanomas resist T-cell therapy through inflammation-induced reversible dedifferentiation. *Nature*, 490(7420):412–416, 2012.

Thomas P Minka. Expectation propagation for approximate bayesian inference. In *Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence*, pages 362–369. Morgan Kaufmann Publishers Inc., 2001.

José C Pinheiro and Douglas M Bates. Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of computational and Graphical Statistics*, 4(1):12–35, 1995.

Adeline Samson, Marc Lavielle, and France Mentré. Extension of the SAEM algorithm to left-censored data in nonlinear mixed-effects model: Application to HIV dynamics model. *Computational Statistics and Data Analysis*, 51(3):1562–1574, 2006.

SAEM: Stochastic Approximation EM

SAEM-MCMC for regular exponential families (Samson et al., 2006):

$$p(y^{obs}, y^{cens}, \zeta; \varrho) = \exp\{\langle \eta_\varrho, \phi(y^{obs}, y^{cens}, \zeta)\rangle - \Phi(\eta_\varrho)\}, \varrho \in \Theta$$

Iteration k of the algorithm:

- Simulation step: simulation of $(y^{cens(k)}, \zeta^{(k)})$ through the simulation of a Markov Chain having $p(y^{cens}, \zeta | y, \hat{\varrho}_{k-1})$ as stationary distribution

- Stochastic approximation: update s_k according to

$$s_k = s_{k-1} + \gamma_k(\phi(y^{obs}, y^{cens(k)}, \zeta^{(k)}) - s_{k-1})$$

(γ_k) is a decreasing sequence: $\sum_{k=1}^{\infty} \gamma_k = \infty$, $\sum_{k=1}^{\infty} \gamma_k^2 < \infty$

- Maximization step: update $\hat{\varrho}_{k-1}$ according to

$$\hat{\varrho}_k = \arg\max_{\varrho \in \Theta} \{\langle \eta_\varrho, s_k \rangle - \Phi(\eta_\varrho)\}$$

EP as approximate Belief Propagation (Heskes et a.l, 2002)

Aim: approximate $p(y|\theta) = \int p(x, y|\theta)dx$ where $p(x, y|\theta) = \prod_{i=1}^{n} \underbrace{p(y_i|x_i, \theta)p(x_i|x_{i-1}, \theta)}_{\psi_i(x_{i-1}, x_i)}$

$$p(x_i, y|\theta) = \underbrace{p(x_i, y_{1:i}|\theta)}_{\approx\tilde{\alpha}_i(x_i)} \overbrace{p(y_{i+1:n}|x_i, y_{1:i}, \theta)}^{\approx\tilde{\beta}_i(x_i)}, \quad y_{i:i'} = (y_i, y_{i+1}, y_{i+2}, ..., y_{i'})$$

$$p(x_{i-1}, x_i, y|\theta) = \underbrace{p(x_{i-1}, y_{1:i-1}|\theta)}_{\approx\tilde{\alpha}_{i-1}(x_{i-1})} \psi_i(x_{i-1}, x_i) \overbrace{p(y_{i+1:n}|x_i, \theta)}^{\approx\tilde{\beta}_i(x_i)}$$

We define $q_i(x_i) \propto \tilde{\alpha}_i(x_i)\tilde{\beta}_i(x_i)$, and $q(x) = \prod_{i=1}^{n} q_i(x_i)$

EP as approximate Belief Propagation (Heskes et a.l, 2002)

Aim: approximate $p(y|\theta) = \int p(x,y|\theta)dx$ where $p(x,y|\theta) = \prod_{i=1}^{n} \underbrace{p(y_i|x_i,\theta)p(x_i|x_{i-1},\theta)}_{\psi_i(x_{i-1},x_i)}$

$$p(x_i, y|\theta) = \underbrace{p(x_i, y_{1:i}|\theta)}_{\approx\tilde{\alpha}_i(x_i)} \overbrace{p(y_{i+1:n}|x_i, y_{1:i}, \theta)}^{\approx\tilde{\beta}_i(x_i)}, \quad y_{i:i'} = (y_i, y_{i+1}, y_{i+2}, ..., y_{i'})$$

$$p(x_{i-1}, x_i, y|\theta) = \underbrace{p(x_{i-1}, y_{1:i-1}|\theta)}_{\approx\tilde{\alpha}_{i-1}(x_{i-1})} \psi_i(x_{i-1}, x_i) \overbrace{p(y_{i+1:n}|x_i, \theta)}^{\approx\tilde{\beta}_i(x_i)}$$

We define $q_i(x_i) \propto \tilde{\alpha}_i(x_i)\tilde{\beta}_i(x_i)$, and $q(x) = \prod_{i=1}^{n} q_i(x_i)$

EP Algorithm:

- Choose $i \in \{1, \ldots, n\} \to$ update for $\tilde{\alpha}_i(x_i)$ and $\tilde{\beta}_{i-1}(x_{i-1})$
- Calculate $\hat{p}(x_{i-1}, x_i) \propto \tilde{\alpha}_{i-1}(x_{i-1}) \underbrace{\psi_i(x_{i-1}, x_i)}_{\approx\tilde{\beta}_{i-1}(x_{i-1})\tilde{\alpha}_i(x_i)} \tilde{\beta}_i(x_i)$
- Find $\underset{\tilde{\alpha}_i(x_i),\ \tilde{\beta}_{i-1}(x_{i-1})}{\arg\min}$ DKL$\left(\hat{p}(x_{i-1}, x_i|y,\theta) \,\middle|\middle|\, q_{i-1}(x_{i-1})q_i(x_i)\right)$
- Iterate until convergence

EP as approximate BP (log-Likelihood approximation)

From estimated α_i, β_i, we deduce $\tilde{L}(\theta) \approx \log p(y|\theta)$ by replacing the $\psi_i(x_{i-1}, x_i)$ by their approximation in $p(y|\theta) = \int \prod_{i=2}^{n} \psi_i(x_{i-1}, x_i) dx_1 \dots dx_n$

$$
\begin{aligned}
\tilde{L}(\theta) &= \log \int \prod_{i=2}^{n} C_i \tilde{\psi}_i(x_{i-1}, x_i)\ dx_1 \dots dx_n \\
&= \log \int \prod_{i=2}^{n} C_i \exp\left\{\beta_{i-1}^T \phi(x_{i-1}) + \alpha_i^T \phi_i(x_i)\right\}\ dx_1 \dots dx_n \\
&= \sum_{i=2}^{n} \log C_i + \sum_{i=2}^{n} \log \int \exp\left\{(\alpha_i + \beta_i)^T \phi_i(x_i)\right\}\ dx_i - \log \int \exp\left\{\alpha_1^T \phi(x_1)\right\}\ dx_1 \\
&= \sum_{i=2}^{n} \log C_i + \sum_{i=2}^{n} \Phi(\alpha_i + \beta_i) - \Phi(\alpha_1)
\end{aligned}
$$

Then, an optimal θ^{opt} is computed as: $\theta^{\text{opt}} = \arg\max_{\theta} \tilde{L}(\theta)$

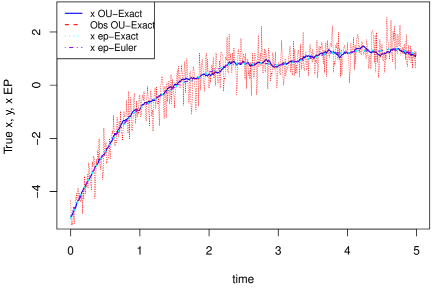
Application 1: Ornstein Uhlenbeck process

$dx_t = \lambda(\mu - x_t)dt + \sigma dW_t, \quad x_{t_0} = x_{\text{initial}}$

Exact solution: $p(x_i|x_{i-1},\theta) = \mathcal{N}\Big(x_i; \mu\left(1 - e^{-\lambda\Delta_i}\right) + x_{i-1}e^{-\lambda\Delta_i}, \frac{\sigma^2}{2\lambda}\left(1 - e^{-2\lambda\Delta_i}\right)\Big)$

Euler discretization: $p(x_i|x_{i-1},\theta) = \mathcal{N}\Big(x_i; x_{i-1} + \lambda(\mu - x_{i-1})\Delta_i, \sigma^2\Delta_i\Big)$

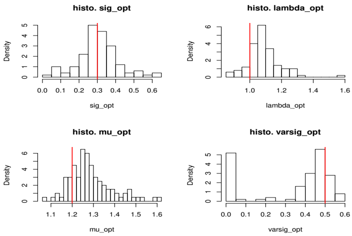
Noisy observations: $y_i = H(x_i) + \epsilon_i = x_i + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \varsigma^2), \quad \theta = (\lambda, \mu, \sigma, \varsigma)$



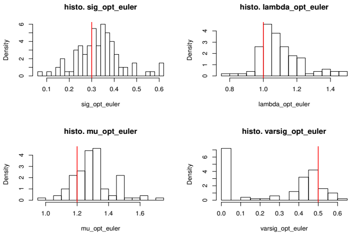
$x_0 = -5, \ \lambda = 1, \ \mu = 1.2, \ \sigma = 0.3, \ \varsigma = 0.5, \ \Delta_i = 10^{-2}, \ n = 500$

Ornstein Uhlenbeck: Parameter estimation with EP

Exact solution



Euler discretization



$$100\ (x, y),\ x_1 = -5,\ \Delta_i = 10^{-2},\ n = 500$$

Application 2: Diffusion approximation of cancer model (In Progress . . .)

Notation: $m_a = \mathbb{E}[x_i|x_{i-1}]$ and $\Sigma_b = \text{Cov}(x_i|x_{i-1})$

$$\tilde{\alpha}_i(x_i) = \mathrm{e}^{\{-\frac{1}{2}x_i^T\alpha_\Lambda x_i + \alpha_{\nu_i}^T x_i\}} \quad \text{and} \quad \tilde{\beta}_i(x_i) = \mathrm{e}^{\{-\frac{1}{2}x_i^T\beta_{\Lambda_i} x_i + \beta_{\nu_i}^T x_i\}}$$

Calculation of two-slice marginal $\hat{p}(x_{i-1}, x_i|y, \theta)$ (under canonical/exponential form)

$$\hat{p}(x_{i-1}, x_i) \propto \tilde{\alpha}_{i-1}(x_{i-1}) \overbrace{\psi_i(x_{i-1}, x_i, y_i, \theta)}^{p(y_i|x_i,\theta)p(x_i|x_{i-1},\theta)} \tilde{\beta}_i(x_i),\ i = 1:n$$

$$= \mathrm{e}^{\{\alpha_{\nu_{i-1}}^T x_{i-1} - \frac{1}{2}x_{i-1}^T \alpha_{\Lambda_{i-1}} x_{i-1}\}} \times \frac{\mathrm{e}^{\{-\dfrac{1}{2}\frac{H(x_i)^2}{\varsigma^2}\}}}{\sqrt{(2\pi)}\,\varsigma} \mathrm{e}^{\{-\frac{1}{2}\frac{y_i{}^2}{\varsigma^2} + y_i\frac{H(x_i)}{\varsigma^2}\}}$$

$$\times \frac{\mathrm{e}^{\{-\dfrac{1}{2}\ m_a^T\Sigma_b^{-1}m_a\}}}{(2\pi)^{d/2}\sqrt{|\det(\Sigma_b)|}} \mathrm{e}^{\{-\frac{1}{2}x_i^T\Sigma_b^{-1}x_i + x_i^T\Sigma_b^{-1}m_a\}} \times \mathrm{e}^{\{\beta_{\nu_i}^T x_i - \frac{1}{2}x_i^T\beta_{\Lambda_i}x_i\}}$$

Application 2: Diffusion approximation of cancer model (In Progress . . .)

Notation: $m_a = \mathbb{E}[x_i|x_{i-1}]$ and $\Sigma_b = \text{Cov}(x_i|x_{i-1})$

$$\tilde{\alpha}_i(x_i) = \mathrm{e}^{\{-\frac{1}{2}x_i^T\alpha_\Lambda x_i + \alpha_{\nu_i}^T x_i\}} \quad \text{and} \quad \tilde{\beta}_i(x_i) = \mathrm{e}^{\{-\frac{1}{2}x_i^T\beta_{\Lambda_i} x_i + \beta_{\nu_i}^T x_i\}}$$

Calculation of two-slice marginal $\hat{p}(x_{i-1}, x_i|y, \theta)$ (under canonical/exponential form)

$$\hat{p}(x_{i-1}, x_i) \propto \tilde{\alpha}_{i-1}(x_{i-1}) \overbrace{\psi_i(x_{i-1}, x_i, y_i, \theta)}^{p(y_i|x_i,\theta)p(x_i|x_{i-1},\theta)} \tilde{\beta}_i(x_i), \; i = 1:n$$

$$= \mathrm{e}^{\{\alpha_{\nu_{i-1}}^T x_{i-1} - \frac{1}{2}x_{i-1}^T\alpha_{\Lambda_{i-1}}x_{i-1}\}} \times \frac{\mathrm{e}^{\{-\frac{1}{2}\frac{H(x_i)^2}{\varsigma^2}\}}}{\sqrt{(2\pi)}\,\varsigma} \mathrm{e}^{\{-\frac{1}{2}\frac{y_i{}^2}{\varsigma^2} + y_i\frac{H(x_i)}{\varsigma^2}\}}$$

$$\times \frac{\mathrm{e}^{\{-\frac{1}{2}\; m_a^T\Sigma_b^{-1}m_a\}}}{(2\pi)^{d/2}\sqrt{|\det(\Sigma_b)|}} \mathrm{e}^{\{-\frac{1}{2}x_i^T\Sigma_b^{-1}x_i + x_i^T\Sigma_b^{-1}m_a\}} \times \mathrm{e}^{\{\beta_{\nu_i}^T x_i - \frac{1}{2}x_i^T\beta_{\Lambda_i}x_i\}}$$

Minimization of DKL$\left(\hat{p}(x_{i-1}, x_i|y, \theta) \;\middle|\middle|\; q_{i-1}(x_{i-1})q_i(x_i)\right)$
(By moment matching: integration and projection on to the chosen family of distribution)

$$q_{i-1}^{\text{new}}(x_{i-1}) \propto \int \hat{p}(x_{i-1}, x_i)dx_i \propto \mathrm{e}^{\{-\frac{1}{2}x_{i-1}^T\Lambda_{i-1}^{\text{new}}x_{i-1} + \nu_{i-1}^{\text{new}}x_{i-1}\}} \rightarrow (\beta_{\nu_{i-1}}^{\text{new}}, \beta_{\Lambda_{i-1}}^{\text{new}})$$

$$q_i^{\text{new}}(x_i) \propto \int \hat{p}(x_{i-1}, x_i)dx_{i-1} \propto \mathrm{e}^{\{-\frac{1}{2}x_i^T\Lambda_i^{\text{new}}x_i + \nu_i^{\text{new}}x_i\}} \rightarrow (\alpha_{\nu_i}^{\text{new}}, \alpha_{\Lambda_i}^{\text{new}})$$

Moment closure approximation for cancer diffusion model

Notation: $x_t^{[l]} \equiv x_{Z_t}$, for $(Z, l) \in \{(M, 1),\ (D, 2),\ (T, 3)\ (A, 4)\}$

Time evolution equation for the first and second order moments:

$$\frac{d\mathbb{E}[x_t^{[l]}]}{dt} = \mathbb{E}[a_l(x_t, \theta)],\ l \in \{1,\ 2,\ 3,\ 4\},$$

$$\frac{d\mathbb{E}[x_t^{[l]} x_t^{[m]}]}{dt} = \mathbb{E}[x_t^{[l]} \times a_m(x_t, \theta)] + \mathbb{E}[x_t^{[m]} \times a_l(x_t, \theta)] + \mathbb{E}[b_{lm}(x_t, \theta)],\ l, m \in \{1,\ 2,\ 3,\ 4\}$$

Two order Moment Approximation (2-MA): Taylor series expansion around $\mu_t = \mathbb{E}[x_t]$ + ignoring terms above degree 2:

$$\frac{d\mu_t^{[l]}}{dt} = a_l(\mu_t, \theta) + \frac{1}{2} \frac{\partial^2 a_l}{\partial x_t \partial x_t^T} : \Sigma_t,$$

$$\frac{d\Sigma_t^{[lm]}}{dt} = \frac{\partial a_l}{\partial x_t^T} \Sigma_t^{[.m]} + \Sigma_t^{[l.]} \frac{\partial a_m}{\partial x_t} + b_{lm}(\mu) + \frac{1}{2} \frac{\partial^2 b_{lm}}{\partial x_t \partial x_t^T} : \Sigma_t,$$

We solve the equations by using Euler scheme:

$$\tilde{\mu}_i^{[l]} = \tilde{\mu}_{i-1}^{[l]} + \Big(a_l(\tilde{\mu}_{i-1}, \theta) + \frac{1}{2} \frac{\partial^2 a_l}{\partial x_t \partial x_t^T} : \tilde{\Sigma}_{i-1}\Big)\Delta_i,$$

$$\tilde{\Sigma}_i^{[lm]} = \tilde{\Sigma}_{i-1}^{[lm]} + \Big(\frac{\partial a_l}{\partial x_t^T} \tilde{\Sigma}_{i-1}^{[.m]} + \tilde{\Sigma}_{i-1}^{[l.]} \frac{\partial a_m}{\partial x_t} + b_{lm}(\tilde{\mu}_{i-1}) + \frac{1}{2} \frac{\partial^2 b_{lm}}{\partial x_t \partial x_t^T} : \tilde{\Sigma}_{i-1}\Big)\Delta_i$$