Entropic Dynamics on Gibbs Statistical Manifolds

Pedro Pessoa

Department of Physics
University at Albany
State University of New York

ppessoa@albany.edu

May 5, 2021

P. Pessoa, F. Xavier Costa, A. Caticha. *Entropy* 23(5), 494 (2021)

MaxEnt

The method of maximum entropy

$$\begin{aligned} \max_{\rho} \quad & S[\rho|q] = -\int dx\ \rho(x) \log\left(\frac{\rho(x)}{q(x)}\right) \ , \\ s.t. \quad & \int dx\ \rho(x) = 1, \\ & \int dx\ \rho(x) a^i(x) = A^i \ . \end{aligned} \tag{1}$$

MaxEnt

The method of maximum entropy

$$
\begin{aligned}
\max_{\rho} \quad & S[\rho|q] = -\int dx\ \rho(x) \log\left(\frac{\rho(x)}{q(x)}\right) \ , \\
s.t. \quad & \int dx\ \rho(x) = 1, \\
& \int dx\ \rho(x) a^i(x) = A^i \ .
\end{aligned} \tag{1}
$$

The result is known Gibbs distributions

$$
\rho(x|\lambda) = \frac{q(x)}{Z(\lambda)} \exp\left(-\sum_{i=1}^{n} \lambda_i a^i(x)\right) \ , \tag{2}
$$

also referred to as exponential family.

- $q(x)$ – prior
- $a^i(x)$ – sufficient statistics
- λ_i – Lagrange multipliers
- $Z(\lambda)$ – Normalizer/ Partition function

Gibbs Distribution

$$\rho(x|\lambda) = \frac{q(x)}{Z(\lambda)} \exp\left(-\sum_{i=1}^{n} \lambda_i a^i(x)\right) \ , \tag{3}$$

Useful properties of the Gibbs distribution:

$$A^i = \left\langle a^i(x) \right\rangle = \frac{\partial F}{\partial \lambda_i} \quad \text{where} \quad F(\lambda) = -\log Z(\lambda) \ . \tag{4}$$

If we calculate the entropy for the Gibbs distribution we have

$$S(A) = -\int dx \ \rho(x|\lambda(A)) \log \frac{\rho(x|\lambda(A))}{q(x)} = \lambda_i(A) A^i - F(\lambda(A)) \ . \tag{5}$$

Gibbs Distribution

$$\rho(x|\lambda) = \frac{q(x)}{Z(\lambda)} \exp\left(-\sum_{i=1}^{n} \lambda_i a^i(x)\right) \tag{6}$$

Distribution	λ parameter	Suff. Stat.	Prior
Exponent Polynomial	$\lambda = \beta$	$a(x) = x^k$	uniform
$\rho(x\|\beta) = \frac{\sqrt[k]{\beta}}{\Gamma(1+1/\beta)} e^{-\beta x^k}$			
Gaussian	$\lambda = \left(-\frac{\mu}{\sigma^2}, \frac{1}{2\sigma^2}\right)$	$a(x) = (x, x^2)$	uniform
$\rho(x\|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$			
Multinomial (k)	$\lambda_i = -\log(\theta_i)$	$a^i = x_i$	$q(x) = \prod_{i=1}^{k} x_i!$
$\rho(x\|\theta) = \frac{n!}{x_1!\ldots x_k!}\theta_1^{x_1}\ldots\theta_k^{x_k}$			
Poisson	$\lambda = -\log m$	$a(x) = x$	$q(x) = 1/x!$
$\rho(x\|m) = \frac{m^x}{x!} e^{-m}$			
Gamma	$\lambda = (\alpha, \beta)$	$a = (\log x, x)$	uniform
$\rho(x\|\alpha,\beta) = \frac{x^{-\alpha} e^{-\beta x}}{\beta^{\alpha-1}\Gamma(1-\alpha)}$			

MaxEnt to Entropic Dynamics

MaxEnt

- Ecology – J Harte. *Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics* Oxford University press (2011)
- Economics – A Golan. *Foundations and Trends(R) in Econometrics* 2, 1 (2008).
- Complex Networks – G Bianconi. *Physical Review E* 79 (3), 036114 (2009).

MaxEnt to Entropic Dynamics

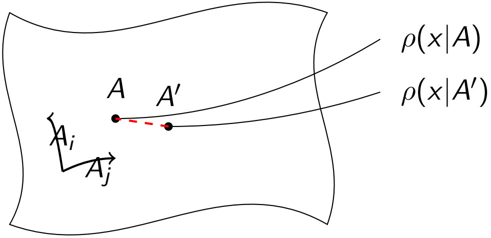
MaxEnt

- Ecology – J Harte. *Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics* Oxford University press (2011)
- Economics – A Golan. *Foundations and Trends(R) in Econometrics* 2, 1 (2008).
- Complex Networks – G Bianconi. *Physical Review E* 79 (3), 036114 (2009).

Entropic Dynamics is a framework in which dynamical laws are obtained by maximizing an entropy. The dynamics is driven by entropy subject to the constraints appropriate to the problem at hand.

- Quantum Mechanics – A Caticha. *Entropy* 21 (10), 943 (2019)
- Quantum Field Theory – S Ipek, M Abedi, A Caticha. *Classical and Quantum Gravity* 36 (20), 205013 (2019)
- Renormalization Groups – P Pessoa, A Caticha. *Entropy* 20 (1), 25 (2018)
- Dynamical models in Economics and Finance – M Abedi, D Bartolomeo. *Entropy* 21 (6), 586 (2019)

Information Geometry



Given the Fisher-Rao metric:

$$g_{ij} = \int dx\ \rho(x|A) \frac{\partial \log \rho(x|A)}{\partial A^i} \frac{\partial \log \rho(x|A)}{\partial A^j}\ . \tag{7}$$

The distance are a measurement of distinguishability, $dl = \sqrt{g_{ij} dA^i dA^j}$.
This metric is UNIQUE – N N Cencov. *Statistical decision rules and optimal inference*, American Mathematical Society (1981)

Information Geometry - Gibbs distributions

The Fisher-Rao metric,

$$g_{ij} = \int dx \ \rho(x|A) \frac{\partial \log \rho(x|A)}{\partial A^i} \frac{\partial \log \rho(x|A)}{\partial A^j} \ , \tag{8}$$

calculated for Gibbs distributions

$$\rho(x|\lambda) = \frac{q(x)}{Z(\lambda)} \exp\left\{ \left(-\lambda_i a^i(x) \right) \right\} \ . \tag{9}$$

- The metric is the covariance matrix.

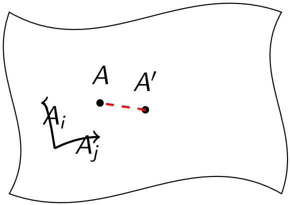
$$g_{ij} = C_{ij} \quad \text{where} \quad C^{ij} = \langle a^i a^j \rangle - A^i A^j = -\frac{\partial A^i}{\partial \lambda_j} \ . \tag{10}$$

- The metric can be represented in the useful forms

$$g_{ij} = -\frac{\partial^2 S}{\partial A^i \partial A^j} \quad \text{and} \quad g^{ij} = -\frac{\partial^2 F}{\partial \lambda_i \partial \lambda_j} \ . \tag{11}$$

Designing a Dynamical System

CHANGE HAPPENS: That change is here labeled by the parameters of the probability distributions (or macrostates) A changes or moves to A'



In the manifold describing the chosen exponential family

$$\rho(x|A) = \frac{q(x)}{Z} \exp\{-\lambda_i(A) a^i(x)\}$$

The Dynamical System is designed by maximizing the joint entropy for both x 'the microstate' and A the expected values or 'macrostates'.

$$\mathrm{S}[P] = -\int dA' \int dx' \ P(x', A'|x, A) \log \left(\frac{P(x', A'|x, A)}{Q(x', A'|x, A)} \right) \quad . \tag{12}$$

Designing a Dynamical System - Entropic Dynamics

- **The prior:** The system moves continuously

$$Q(x', A'|x, A) = Q(x'|x)Q(A'|A) = q(x')Q(A'|A) \tag{13}$$

$$Q(A'|A) \propto g^{1/2}(A') \exp\left[-\frac{1}{2\tau} g_{ij} \Delta A^i \Delta A^j\right] , \tag{14}$$

where $\Delta A = A' - A$ and $g^{1/2} = \det g_{ij}$.

For a small steps implementation we have $\tau \to 0$.

- **The Constraint**

$$P(x', A'|x, A) = P(x'|A', x, A)P(A'|A, x) = \rho(x'|A')P(A'|A) . \tag{15}$$

That means the probability for x' will be a Gibbs distribution, therefore a point of the manifold equivalent to A'. This constraint enforces the idea that the motion is confined to the statistical manifold.

The Dynamics

Maximizing the Entropy under such constraints we obtain:

$$P(A'|A) = \frac{1}{\xi} g^{1/2}(A') \exp\left[S(A) + \frac{\partial S}{\partial A^i}\Delta A^i - \frac{1}{2\tau} g_{ij}\Delta A^i \Delta A^j\right] \ , \qquad (16)$$

We identify entropic time so that motion looks simple

$$\tau = \eta \Delta t \Longrightarrow \Delta t = \frac{g_{ij}}{\eta}\left\langle \Delta A^i \Delta A^j \right\rangle \qquad (17)$$

We can write the transition as a differential equation:

$$P_{t'}(A) = \int dA \ P_{\Delta t}(A'|A) P_t(A) \Longrightarrow \frac{\partial}{\partial t} P = -\frac{\partial}{\partial A^i}\left(P v^i\right) \ . \qquad (18)$$

where the current velocity is

$$v^i = \underbrace{\eta g^{ij} \frac{\partial S}{\partial A^j}}_{\text{Entropic Drift}} - \underbrace{\frac{\eta}{2} g^{ij} \frac{\partial}{\partial A^j} \log\left(\frac{P}{g^{1/2}}\right)}_{\text{``Osmotic'' term}} \ . \qquad (19)$$

The result is a diffusion equation on a curved space.

Examples - 3 state system

For a 3 state system, $\mathcal{X} = 1, 2, 3$ any distribution can be put into the exponential form with coordinates given by

$$A^1 = p(x = 1) = \langle \delta_1^x \rangle, \quad A^2 = p(x = 2) = \langle \delta_2^x \rangle \tag{20}$$

leading to an entropy

$$S = -\sum_{i=1}^{3} \rho(i) \log(\rho(i)) = -\sum_{i=1}^{3} A^i \log\left(A^i\right) \tag{21}$$

and metric

$$g_{ij} = \begin{bmatrix} \frac{1}{A^3} + \frac{1}{A^1} & \frac{1}{A^3} \\ \frac{1}{A^3} & \frac{1}{A^3} + \frac{1}{A^2} \end{bmatrix}, \tag{22}$$

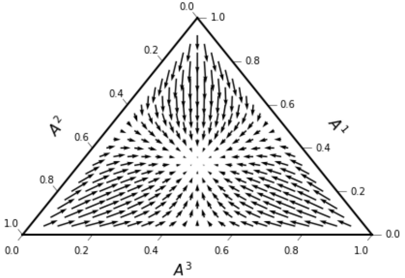
Examples - 3 state system

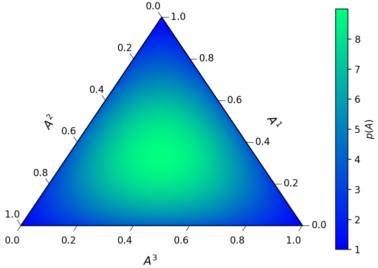
Drift velocity

$$v_d^1 = A^1 \left[A^2 \log\left(\frac{A^2}{A^3}\right) + (A^1 - 1)\log\left(\frac{A^1}{A^3}\right)\right]$$
$$v_d^2 = A^2 \left[A^1 \log\left(\frac{A^1}{A^3}\right) + (A^2 - 1)\log\left(\frac{A^2}{A^3}\right)\right] \quad (23)$$

Static probability is

$$P(A) \propto g^{1/2} \prod_{i=1}^{3} \left(A^i\right)^{-2A^i} \quad (24)$$





P. Pessoa, F. Xavier Costa, A. Caticha. *Entropy* 23(5), 494 (2021)

Conclusion

- We were able to create a dynamics for the values of A, taking into account the natural geometric structure.
- This dynamical system is obtained as a form of entropic inference and it can be applied to any form of Gibbs/exponential distributions (the microstates can be anything).
- It depends only on the calculation of S and g_{ij} is straightforward, albeit laborious, by standard methods and have a clear statistical interpretation.

Further work

- F Xavier Costa, P Pessoa.
  Entropic Dynamics of Networks
  Northeast J. Complex Syst. 3(1), 4 (2021)
- P Pessoa.
  Legendre transformation and information geometry for the maximum entropy theory of ecology
  Under review Preprint: https://arxiv.org/abs/2103.11230

Thank You

Work developed with F. Xavier Costa and A.Caticha. Special thanks to N. Caticha, C. Cafaro, S. Ipek, N. Carrara, and M. Abedi for important discussions and questions in the development of this article.

OBRIGADO

Thank You

Work developed with F. Xavier Costa and A.Caticha. Special thanks to N. Caticha, C. Cafaro, S. Ipek, N. Carrara, and M. Abedi for important discussions and questions in the development of this article.

OBRIGADO

References

- P. Pessoa, F. Xavier Costa, A. Caticha. *Entropy* 23(5), 494 (2021)
- J Harte. *Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics*, Oxford University Press (2011)
- A Golan. *Foundations and Trends(R) in Econometrics* 2, 1 (2008).
- G Bianconi. *Physical Review E* 79 (3), 036114 (2009).
- A Caticha. *Entropy* 21 (10), 943 (2019)
- S Ipek, M Abedi, A Caticha. *Classical and Quantum Gravity* 36 (20), 205013 (2019)
- M Abedi, D Bartolomeo. *Entropy* 21 (6), 586 (2019)
- P Pessoa, A Caticha. *Entropy* 20 (1), 25 (2018)
- N N Cencov. *Statistical decision rules and optimal inference*, American Mathematical Society (1981)
- F Xavier Costa, P Pessoa. *Northeast J. Complex Syst.* 3(1), 4 (2021)
- P Pessoa. https://arxiv.org/abs/2103.11230