
Modèle dynamique pour la représentation de données d'expression des gènes en cellules uniques

Mathilde Gaillard^{*†1} and Ulysse Herbach^{‡1}

¹Inria – L'Institut National de Recherche en Informatique et en Automatique (INRIA) – France

Résumé

Une première simplification du mécanisme d'expression des gènes considère qu'un gène est transcrit en ARN messenger, lui-même traduit à son tour en protéine. Les données issues de cellules individuelles (single-cell) ont permis de révéler la présence d'une variabilité biologique entre des cellules de génome et environnement identiques, mettant en lumière non seulement les aspects épigénétiques mais également le caractère stochastique de l'expression des gènes. Cette stochasticité soulève deux enjeux d'un point de vue statistique. Tout d'abord, il faut pouvoir estimer la variabilité entre les cellules (ou entre les gènes) en utilisant par exemple l'entropie de Shannon (Gandrillon et al., 2021). D'autre part, dans le contexte des réseaux de régulation qui sous-tendent les états et types cellulaires, il s'agit de construire un modèle qui tienne compte à la fois de la stochasticité et de l'interaction des gènes entre eux. On s'intéresse ici à un modèle dynamique d'expression des gènes, formulé comme un processus de Markov déterministe par morceaux (PDMP) et décrivant un nombre arbitraire de gènes en interaction (Herbach et al., 2017). Ce modèle stochastique reproduit la variabilité biologique mesurée expérimentalement (Ventre et al., 2023), tout en étant compatible avec l'observation d'un pic transitoire dans l'entropie de Shannon au cours de processus de différenciation comme l'hématopoïèse (Dussiau et al., 2022). Nous justifions d'abord l'existence et l'unicité d'une loi invariante pour ce modèle et estimons la vitesse de convergence vers cette dernière. Cette loi admet alors une expression analytique pour toute une classe de paramètres et peut s'interpréter comme un modèle graphique probabiliste de type champ de Markov caché, avec de bonnes propriétés pour l'inférence des interactions.

Mots-Clés: Processus de Markov déterministes par morceaux, Distance de Wasserstein, Champs aléatoires de Markov

*Intervenant

†Auteur correspondant: mathilde.gaillard@inria.fr

‡Auteur correspondant: ulysse.herbach@inria.fr