Length of stay prediction using Factor Graphs

Johann & Jannis

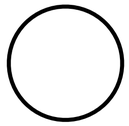
2023, Dec. 05

Factor Graphs

very brief introduction

- graphical probabilistic model
- model factor structure of joint probabilities as a bipartite graph of

factors and



variables

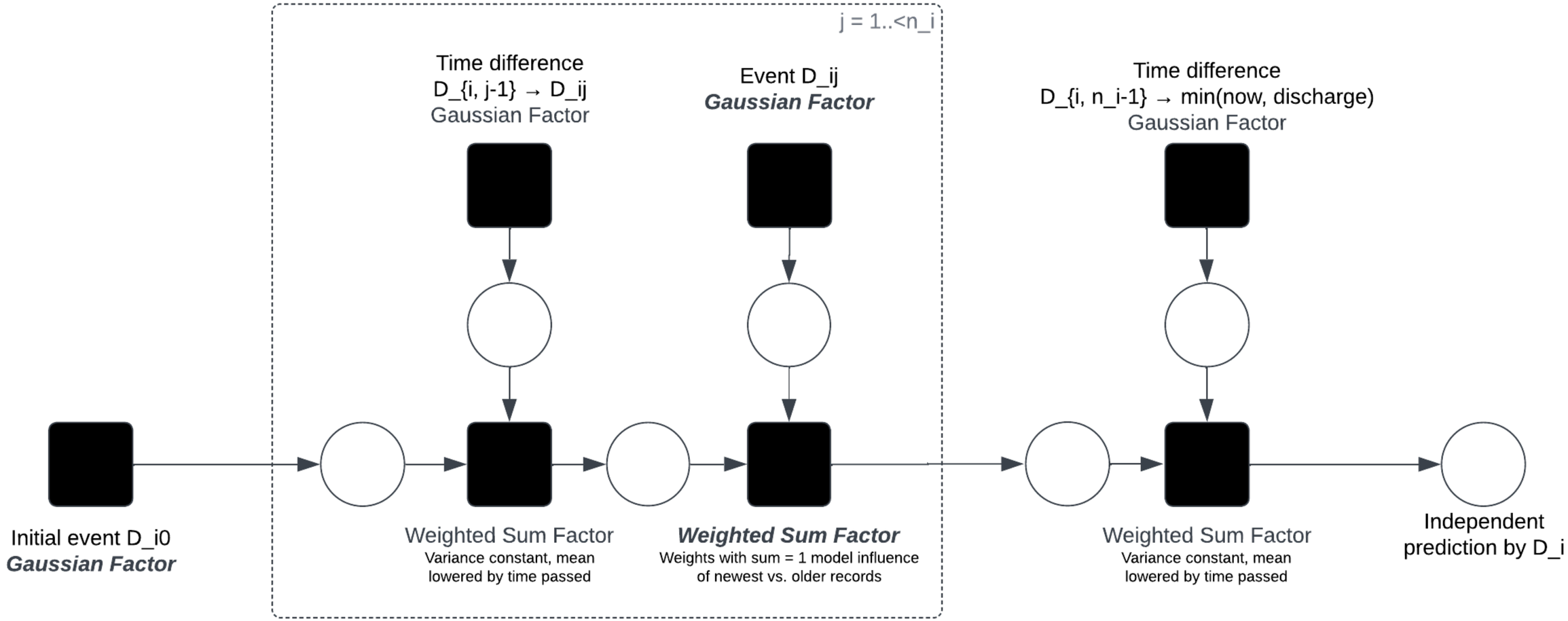
- example application: Bayesian Ranking, e.g. which ad do you see on Amazon
- further reading: Frey, Brendan J, Hans-Andrea Loeliger, Frank R Kschischang, and Niclas Wiberg. 1998 “Factor Graphs and Algorithms.” or watching: https://www.tele-task.de/lecture/video/10061/ and following

Factor Graphs

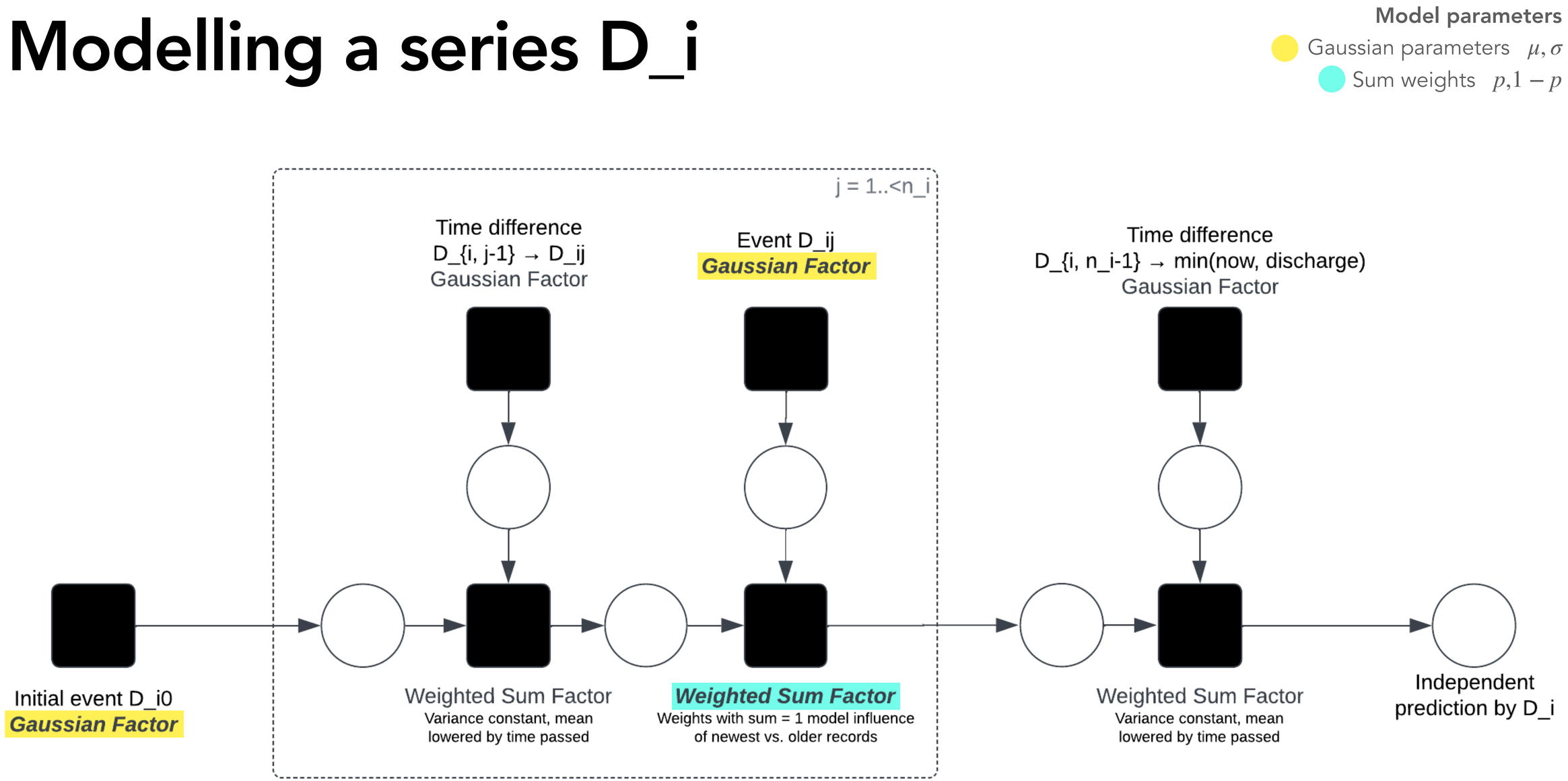
length of stay model & assumptions

- during a stay, multiple series of temporal data are recorded
  - e.g. daily temperature measurements as one series
- the most recent data point in each series is the most relevant
- new records have additive influence on length of stay
- outcome of record is normally distributed
- records in different series are independent

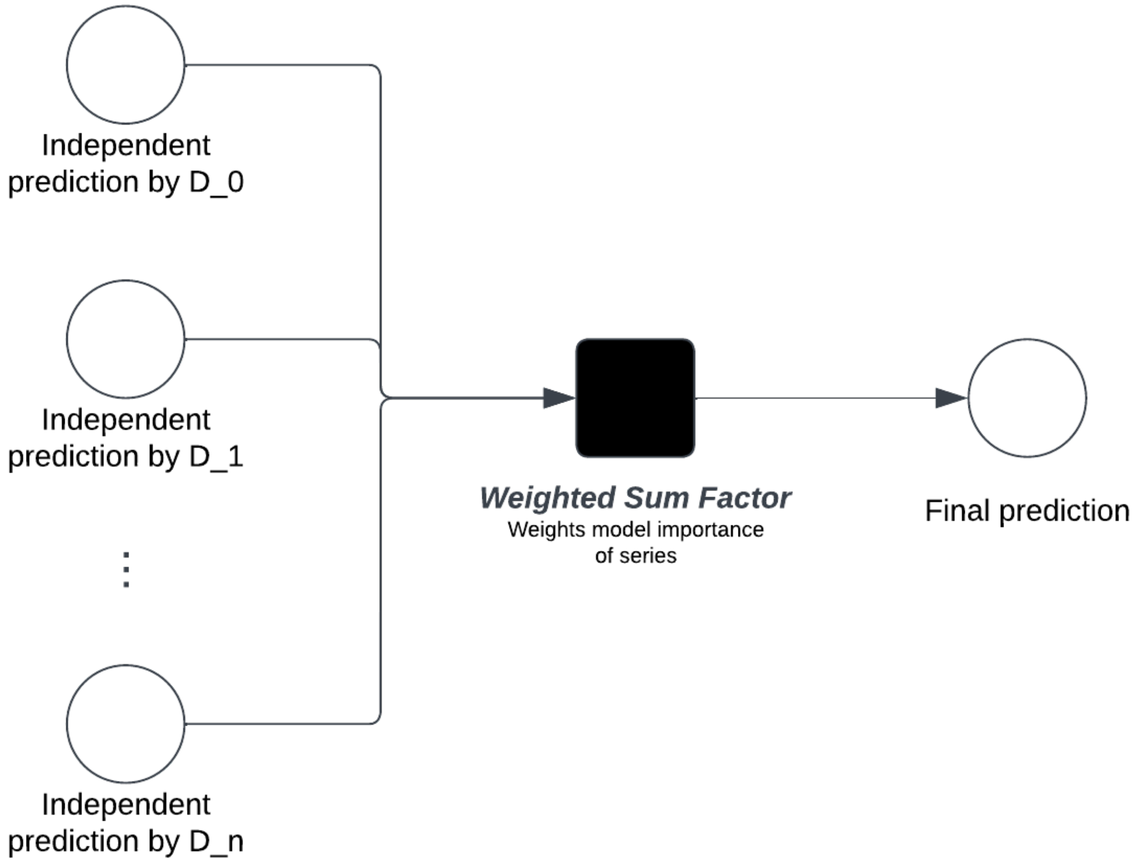
Modelling a series D_i



Modelling a series D_i



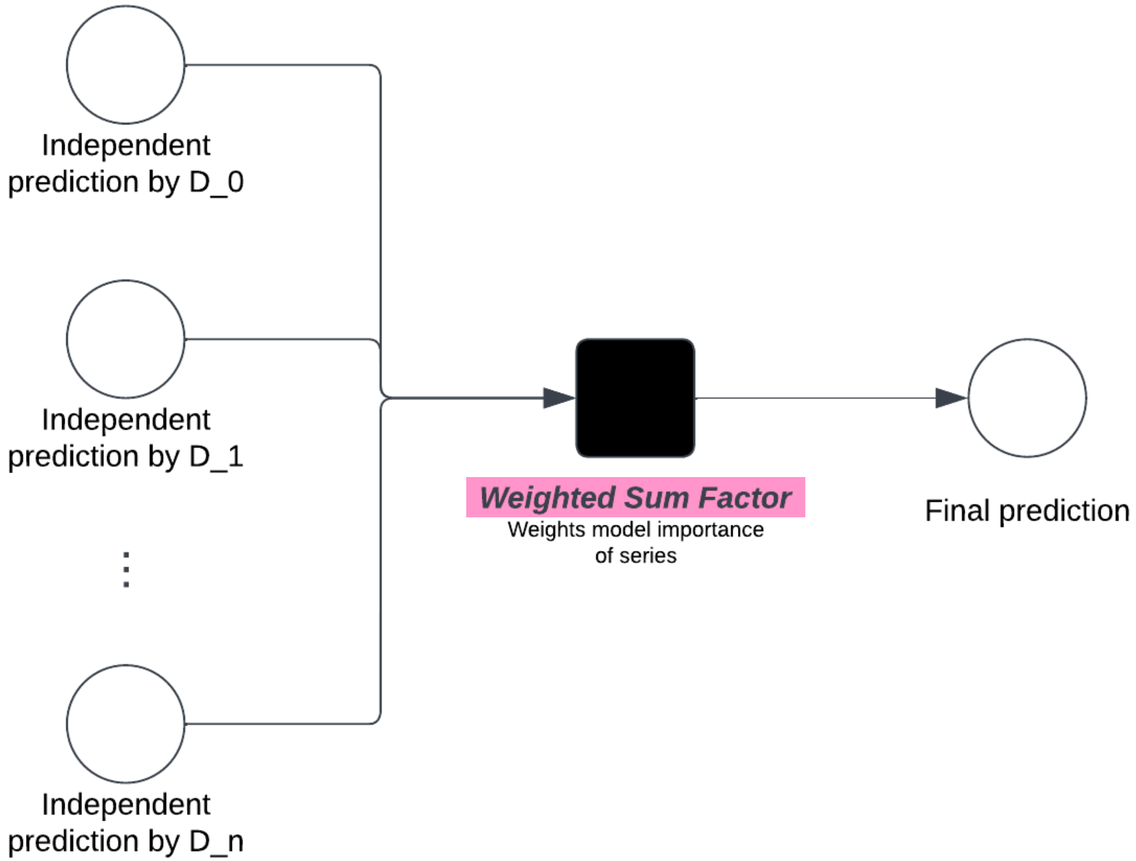
Final prediction



Final prediction

Model parameters

Sum weights $p_0, p_1, \ldots, p_i$ scaled to have sum of 1



Training

Parameters

- Sum weights $p, p-1$ for exponential rolling average in each series, i.e. “How important is a new record compared to all old ones?”
- Gaussian parameters for all records, i.e. “How does this record influence the mean & variance of length of stay”
  - for continuous records like temperature: map value to parameters
- Weights of final sum for each series, i.e. “How influential is this series overall”

Training

Process

- Build factor graphs for
  - all stays
  - subsets of records for each stay by omitting $1 \ldots n-1$ last (ordered by time) records
- use gradient descent to find maximum likelihood parameters

Outcomes

- very explainable & easy to understand model
- a normal distribution is modeled for each record
  - very interesting, e.g. to compare our diagnosis distributions with catalog stays and validate model or to find out exactly what a record implies for the length of stay
- weighted sum weights also have interesting meaning