Why is the self transition rate qii negative lambda?

question posted 21 days ago by **gmikawa**

Is the definition of qii negative lambda? I would like to know the intuitive meaning of the qii.

This post is visible to everyone.

Add a Response

2 responses

laurent_decreusefond(Staff)
21 days ago

There is no intuitive meaning of the minus sign. Say it is a mathematical convention.

Be careful that it is not a self transition rate, the absolute value of q_{ii} is the parameter of the exponential distribution which defines the time period during which the process stays in state i before jumping to another state.

Thank you, Laurent. Is there any specific name for qii? qij is transition rate from i to j, qii is transition rate from i to i. So, I thought the transition from i to i was self-transition. Self-transition in CTMC does not make any sense? It cannot be observed anyway unless it is transitioned, can it?

posted 21 days ago by**gmikawa**

Add a comment

Sandrine_VATON (Staff)
21 days ago

Let us define q_{ii} as $q_{ii} = -\sum_{j \neq i} q_{ij}$. This quantity is negative since the quantities q_{ij}, when $j \neq i$ have been defined as transition rates (so, they have positive values).

It holds that:

$$\pi_j(t + \Delta t) = \sum_{i \neq j} \pi_i(t) q_{ij} \Delta t + \pi_j(t)(1 + q_{jj} \Delta t) + o(\Delta t)$$

From this equation we obtain that:

$$\frac{d}{dt} \pi_j(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} [\pi_j(t + \Delta t) - \pi_j(t)] = \sum_i \pi_i(t) q_{ij}$$

And consequently:

$$\pi(t) = \pi(0) \exp(Qt)$$

where the state probabilities are written as a row vector and where the exponential is a matrix exponential.

We can remark that the element (i,j) of the matrix exponential is a conditional probability:

$$P(X(t + \Delta t) = j \mid X(t) = i)$$

so that it should sum to 1 for any value of t.

Now, consider a first order Taylor expansion of the matrix exponential:

$$\exp(Qt) = I + Qt + o(t)$$

As the sum of the elements on each row of the upper matrix should be equal to 1 whatever the value of t it comes that the sum of the elements on each row of the matrix Q should be equal to 0.

You can also consider as mentioned by Laurent that the diagonal element being equal to the opposite of the other elements on the corresponding row is a mathematical convention.

Thank you, Sandrine. The second term in the right in the first equation, I thought it was π(i)(t)q(jj)Δt. May I ask you how the second term becomes the one in the right in the first equation?

posted 21 days ago by**gmikawa**

The term $(1 + q_{jj}\Delta t)$ corresponds to the probability that the system remains in state j during the interval $[t, t + \Delta t]$. The probability that the system leaves state j would be $\sum_{i \neq j} q_{ji} \Delta t = -q_{jj} \Delta t$
.

posted 21 days ago by**Sandrine_VATON** (Staff)

Thank you, Sandrine. That helps a lot. It would be nice if you could tell me the rational reason for the definition of qii=-Σqij, the sum of over j except i.

posted 21 days ago by**gmikawa**

I think Sandrine's last comment answers your question gmikawa, right?

I can try an alternative formulation:

- Assume we are in state i. Over a small duration Δ_t, some event can occur and make us switch to another state j(the probability that more than one event occurs during Δ_t is a $o(\Delta_t)$, if you're not convinced about that we can come back to that).
- What are those events, and when can they occur? Well, the event making us go to State j is characterized by the rate q_{ij}, which means it should occur after some exponentially distributed time X with parameter q_{ij}: the probability that this event does **not** occur in the next Δ_t seconds is $\mathbf{P}(X > \Delta_t) = e^{-q_{ij}\Delta_t}$
- So over Δ_t, we stay in the same state if none of those events occur within Δ_t. But all those events are independent, hence the probability that none occurs equals $\Pi_{j\neq i} e^{-q_{ij}\Delta_t} = e^{-\sum_{j\neq i} q_{ij}\Delta_t} = 1 + q_{ii}\Delta_t + o(\Delta_t)$ as Sandrine wrote

wrote previously. This is the probability that we stay in State i for the next (small) Δ_t seconds.

I'm not sure whether this answers your request of a "rational reason" for considering that sum.

You were also right in your comment before: self-transition in continuous-time does not make any sense (staying in the current state is the "default" behavior, we see only transitions to other states)

posted 20 days ago by**PatrickMaille** (Staff)

May I ask you one more question? The transition rate from i to j is defined by qij=λ(i)pij. Is this strictly defined only when i and j are different? Or can I still apply this definition when i = j when pii is restricted to zero?

posted 20 days ago by**gmikawa**

Consider a first order Taylor expansion of the matrix exponential: $exp(Qt) = I + Qt + o(t)$
.

As the sum of the elements on each row of the upper matrix should be equal to 1 whatever the value of t it comes that the sum of the elements on each row of the matrix Q should be equal to 0.

The sum of the elements of row i is equal to 0 if

$$\sum_j q_{ij} = 0$$

which is equivalent to:

$$q_{ii} + \sum_{j \neq i} q_{ij} = 0$$

so that

$$q_{ii} = -\sum_{j \neq i} q_{ij}$$

posted 20 days ago by**Sandrine_VATON** (Staff)

As you said before, in continuous time there is not much sense in talking about a transition rate from a state to itself (staying in the same state is the default behavior over an infinitely small amount of time). So yes, you can apply the formula only for $i \neq j$. In any case, the p_{ij} you are mentioning refer to the transition probabilities *when we leave state* i, so we have $p_{ii} = 0$.

posted 20 days ago by **PatrickMaille** (Staff)

This confused me a bit too since in the discrete case, we used p_{ij} as the transition probabilities. Here it seems p_{ij} is the probability of going from state i to state j given either a λ or μ event happens. Is that right? I.e. in the case of a birth-death process, we're always going to have $p_{i,i+1} = 1$ or $p_{i,i-1} = 1$, why is why the p_{ij}'s don't appear anywhere in the transition matrix even though we previously defined $q_{ij} = \lambda_i p_{ij}$. Instead we just have $p_{i,i+1} = \lambda_i$ and $p_{i,i-1} = \mu$.

posted 20 days ago by **alexanderwiguna**

Thank you, Patrick. I love the way you showed me the remaining probability in the state j. Then, combining the remaining probability with Sandrine's comment, the departure probability is 1-(1+piiΔt)=-piiΔt=λiΔt, which is indeed the transition out of state i. Can I interpret λiΔt as transition probability out of state i in Δt?

posted 20 days ago by**gmikawa**

Sandrine, could you explain how you solved the differential equation d/dt(πi)=.... and represented it as matrix format?

posted 20 days ago by**gmikawa**

We have solved a system of differential equations. Note that

$$\pi(t) = [\pi_0(t)\pi_1(t)\pi_2(t)\ldots]$$

is a row vector, each element of this vector is a state probability.

With this notation it holds that

$$\frac{d}{dt}\pi(t) = \pi(t)Q$$

The solution of such a system is of matrix exponential form:

$$\pi(t) = \pi(0)exp(Qt)$$

where

$$exp(Qt) = I + Qt + Q^2\frac{t^2}{2} + Q^3\frac{t^3}{3!} + \ldots$$

posted 20 days ago by**Sandrine_VATON** (Staff)

Thank you, Sandrine. From d/dt(π)=πQ, m differential equations are obtained, i.e. d/dt(π1)=π1.q11+...+πm.qm1, ...d/dt(πm)=π1.q1m+...+πm.qmm, when the system has m states. Right? d/dt(π1+...+πm)=0, π1+...+πm=1. Just end up with normalization condition. What I would like to know is how π1=π1(0).exp(q11.t)+...+πm(0).exp(qm1.t), etc are delivered from above m differential equations. When the situation is one dimensional it is easy, but with the m-dimensional I need some help. I am understanding Q is the transition rate matrix with ij element qij and exp(Qt) is the matrix with ij element exp(qij.t).

Blockquote

posted 19 days ago by**gmikawa**

Hi gmikawa,

$exp(Qt)$ is NOT a matrix with elements $exp(q_{ij}t)$. A matrix exponential is defined as the sum of a Taylor series as explained above.

posted 19 days ago by**Sandrine_VATON** (Staff)

It would be helpful if you could write the equation π1 for example instead of using matrix notation. It would also be nice if you could tell me the background of the definition of matrix exponential expanded by Taylor series.

posted 19 days ago by**gmikawa**

Hi gmikawa, what do you call $\pi 1$: the probability of State 1 (at what time?) or the probability at time 1(but for what state?)

About matrix exponentials, the necessary background is just matrix multiplication, then as Sandrine said we can see the exponential of some matrix M as defined by $e^M = \sum_{i=0}^{\infty} \frac{M^i}{i!}$. You can note that it corresponds to the classical exponential function if M is a number (i.e., a 1×1matrix).

posted 19 days ago by**PatrickMaille** (Staff)

Hi Patrick. π1is π1(t), state 1 at time t. I am curious how the metrics π(0)e^(Qt) looks like as an element-wise expression π1(t)={π(0)e^(Qt)}11.

posted 18 days ago by**gmikawa**

I'm not sure we can write it in an easier way than what you just wrote.

Probably if you can diagonalize Q under the form $Q = LDL^{-1}$ with Lan invertible matrix and D a diagonal matrix (with the eigenvalues $v_1, \ldots, v_N$ of Q in the diagonal, with $N$the number of states), then you can have a different writing of the exponential of Qt as $e^{Qt} = LD_tL^{-1}$ with D_t the diagonal matrix with terms $e^{tv_1}, \ldots, e^{tv_N}$. But that would imply finding the eigenvalues of Q.

posted 18 days ago by**PatrickMaille** (Staff)

I try to give you an intuitive explanation to convince you that it is reasonable to have negative values on the diagonal of the matrix Q.

Consider the load balance equations in continuous time. If we write these equations in matrix form we get $\pi Q = 0$

Equivalently we can write the system of load balance equations as $\sum_i \pi_i q_{ij} = 0$ for

$\sum_i \pi_i q_{ij} = 0$ for all j.

Now, we are going to move the term $\pi_j q_{jj}$ to the other side of the equal sign:
$\sum_{i \neq j} \pi_i q_{ij} = -\pi_j q_{jj}$

The terms q_{ij} for $i \neq j$ are transition rates, they have positive values. As you can see from the upper equation, the diagonal term q_{jj} should consequently have a negative value.

Note that q_{jj} is not a transition rate, but the opposite of the parameter λ_j of the exponential random variables that represents the sojourn time in state j. So, $q_{jj} = -\lambda_j$ where the CTMC remains in state j during a time which is distributed as $Exp(\lambda_j)$.

posted 17 days ago by **Sandrine_VATON** (Staff)

Thank you, Sandrine. I like your intuitive explanation. qij is transition rate, but qjj is not. How come do you apply the same notation for a different event? How should I call qjj?

posted 17 days ago by **gmikawa**

Yes, exactly!

q_{jj} is not a transition rate, but the opposite of the parameter of the exponential sojourn time λ_j in state j.

We apply the "same" notation because q_{ij} and q_{jj} are two elements of a same matrix Q. And there is no particular word for the diagonal term q_{jj}.

Note that in the course Q is called the "transition rate matrix". Another probably more usual name for that matrix is "infinitesimal generator". We have opted for "transition rate matrix" because we think that this is more intuitive than "infinitesimal generator".

But you are right, the diagonal elements are not transition rates! And this was maybe misleading.

posted 16 days ago by**Sandrine_VATON** (Staff)

I appreciate your clarification. I agree that transition rate matrix more sounds more intuitive, as I can grasp immediately what Q does.

posted 16 days ago by**gmikawa**