Stochastic Matrices and the Perron–Frobenius Theorem*

René Thiemann

March 17, 2025

Abstract

Stochastic matrices are a convenient way to model discrete-time and finite state Markov chains. The Perron–Frobenius theorem tells us something about the existence and uniqueness of non-negative eigenvectors of a stochastic matrix.

In this entry, we formalize stochastic matrices, link the formalization to the existing AFP-entry on Markov chains, and apply the Perron–Frobenius theorem to prove that stationary distributions always exist, and they are unique if the stochastic matrix is irreducible.

Contents

1 Introduction 1

2 Stochastic Matrices 2

3 Stochastic Vectors and Probability Mass Functions 3

4 Stochastic Matrices and Markov Models 5

5 Eigenspaces 6

6 Stochastic Matrices and the Perron–Frobenius Theorem 7

1 Introduction

In their AFP entry Markov Models [2], Hölzl and Nipkow provide a framework for specifying discrete- and continuous-time Markov chains.

In the following, we instantiate their framework by formalizing right-stochastic matrices and stochastic vectors. These vectors encode probability

*Supported by FWF (Austrian Science Fund) project Y757.

mass functions over a finite set of states, whereas stochastic matrices can be utilized to model discrete-time and finite space Markov chains.

The formulation of Markov chains as matrices has the advantage that certain concepts can easily be expressed via matrices. For instance, a stationary distribution is nothing else than a non-negative real eigenvector of the transition matrix for eigenvalue 1. As a consequence, we can derive certain properties on Markov chains using results on matrices. To be more precise, we utilize the formalization of the Perron–Frobenius theorem [1] to prove that a stationary distribution always exists, and that it is unique if the transition matrix is irreducible.

2 Stochastic Matrices

We define a type for stochastic vectors and right-stochastic matrices, i.e., non-negative real vectors and matrices where the sum of each column is 1. For this type we define a matrix-vector multplication, i.e., we show that $A*v$ is a stochastic vector, if A is a right-stochastic matrix and v a stochastic vector.

theory *Stochastic-Matrix*
  imports *Perron-Frobenius.Perron-Frobenius-Aux*
begin

definition *non-neg-vec* :: *'a* :: *linordered-idom* ^ *'n* ⇒ *bool* **where**
  non-neg-vec A ≡ (∀ *i*. *A* $ *i* ≥ *0*)

definition *stoch-vec* :: *'a* :: *comm-ring-1* ^ *'n* ⇒ *bool* **where**
  stoch-vec v = (*sum* (λ *i*. *v* $ *i*) *UNIV* = *1*)

definition *right-stoch-mat* :: *'a* :: *comm-ring-1* ^ *'n* ^ *'m* ⇒ *bool* **where**
  right-stoch-mat a = (∀ *j*. *stoch-vec* (*column j a*))

typedef *'i st-mat* = { *a* :: *real* ^ *'i* ^ *'i*. *non-neg-mat a* ∧ *right-stoch-mat a*}
  morphisms *st-mat Abs-st-mat*
  ⟨*proof*⟩

setup-lifting *type-definition-st-mat*

typedef *'i st-vec* = { *v* :: *real* ^ *'i*. *non-neg-vec v* ∧ *stoch-vec v*}
  morphisms *st-vec Abs-st-vec*
  ⟨*proof*⟩

setup-lifting *type-definition-st-vec*

lift-definition *transition-vec-of-st-mat* :: *'i* :: *finite st-mat* ⇒ *'i* ⇒ *'i st-vec*
  is λ *a i*. *column i a*
  ⟨*proof*⟩

lemma *non-neg-vec-st-vec*: *non-neg-vec* (*st-vec v*)
⟨*proof*⟩

lemma *non-neg-mat-mult-non-neg-vec*: *non-neg-mat a* $\Longrightarrow$ *non-neg-vec v* $\Longrightarrow$
non-neg-vec (*a* ∗*v v*)
⟨*proof*⟩

lemma *right-stoch-mat-mult-stoch-vec*: **assumes** *right-stoch-mat a* **and** *stoch-vec v*
shows *stoch-vec* (*a* ∗*v v*)
⟨*proof*⟩

lift-definition *st-mat-times-st-vec* :: *'i* :: *finite st-mat* ⇒ *'i st-vec* ⇒ *'i st-vec*
(**infixl** ‹∗*st*› *70*) **is** (∗*v*)
⟨*proof*⟩

lift-definition *to-st-vec* :: *real* ^ *'i* ⇒ *'i st-vec* **is**
λ *x*. *if stoch-vec x* ∧ *non-neg-vec x then x else* (χ *i*. *if i* = *undefined then 1 else 0*)
⟨*proof*⟩

lemma *right-stoch-mat-st-mat*: *right-stoch-mat* (*st-mat A*)
⟨*proof*⟩

lemma *non-neg-mat-st-mat*: *non-neg-mat* (*st-mat A*)
⟨*proof*⟩

lemma *st-mat-mult-st-vec*: *st-mat A* ∗*v st-vec X* = *st-vec* (*A* ∗*st X*) ⟨*proof*⟩

lemma *st-vec-nonneg*[*simp*]: *st-vec x* $ *i* ≥ *0*
⟨*proof*⟩

lemma *st-mat-nonneg*[*simp*]: *st-mat x* $ *i* $*h j* ≥ *0*
⟨*proof*⟩

end

3 Stochastic Vectors and Probability Mass Functions

We prove that over a finite type, stochastic vectors and probability mass functions are essentially the same thing: one can convert between both representations.

theory *Stochastic-Vector-PMF*
imports *Stochastic-Matrix*
HOL−Probability.Probability-Mass-Function
begin

lemma *sigma-algebra-UNIV-finite*[*simp*]: *sigma-algebra* (*UNIV* :: *'a* :: *finite set*) *UNIV*
⟨*proof*⟩

definition *measure-of-st-vec'* :: *'a st-vec* ⇒ *'a* :: *finite set* ⇒ *ennreal* **where**
measure-of-st-vec' x I = *sum* (λ *i*. *st-vec x* $ *i*) *I*

lemma *positive-measure-of-st-vec'*[*simp*]: *positive A* (*measure-of-st-vec' x*)
⟨*proof*⟩

lemma *measure-space-measure-of-st-vec'*: *measure-space UNIV UNIV* (*measure-of-st-vec' x*)
⟨*proof*⟩

context begin
setup-lifting *type-definition-measure*

lift-definition *measure-of-st-vec* :: *'a st-vec* ⇒ *'a* :: *finite measure* **is**
λ *x*. (*UNIV*, *UNIV*, *measure-of-st-vec' x*)
⟨*proof*⟩

lemma *sets-measure-of-st-vec*[*simp*]: *sets* (*measure-of-st-vec x*) = *UNIV*
⟨*proof*⟩

lemma *space-measure-of-st-vec*[*simp*]: *space* (*measure-of-st-vec x*) = *UNIV*
⟨*proof*⟩

lemma *emeasure-measure-of-st-vec*[*simp*]: *emeasure* (*measure-of-st-vec x*) *I* =
sum (λ *i*. *st-vec x* $ *i*) *I*
⟨*proof*⟩

lemma *prob-space-measure-of-st-vec*: *prob-space* (*measure-of-st-vec x*)
⟨*proof*⟩
end

lift-definition *st-vec-of-pmf* :: *'i* :: *finite pmf* ⇒ *'i st-vec* **is**
λ *pmF*. *vec-lambda* (*pmf pmF*)
⟨*proof*⟩

context *pmf-as-measure*
begin
lift-definition *pmf-of-st-vec* :: *'a* :: *finite st-vec* ⇒ *'a pmf* **is** *measure-of-st-vec*
⟨*proof*⟩

lemma *st-vec-st-vec-of-pmf*[*simp*]:
st-vec (*st-vec-of-pmf x*) $ *i* = *pmf x i*
⟨*proof*⟩

lemma *pmf-pmf-of-st-vec*[*simp*]: *pmf* (*pmf-of-st-vec x*) *i* = *st-vec x* $ *i*
⟨*proof*⟩

lemma *st-vec-of-pmf-pmf-of-st-vec*[*simp*]: *st-vec-of-pmf* (*pmf-of-st-vec x*) = *x*
⟨*proof*⟩

lemma *pmf-of-st-vec-inj*: (*pmf-of-st-vec x* = *pmf-of-st-vec y*) = (*x* = *y*)
⟨*proof*⟩
end
end

4 Stochastic Matrices and Markov Models

We interpret stochastic matrices as Markov chain with discrete time and finite state and prove that the bind-operation on probability mass functions is precisely matrix-vector multiplication. As a consequence, the notion of stationary distribution is equivalent to being an eigenvector with eigenvalue 1.

theory *Stochastic-Matrix-Markov-Models*
imports
Markov-Models.Classifying-Markov-Chain-States
Stochastic-Vector-PMF
begin

definition *transition-of-st-mat* :: *'i st-mat* ⇒ *'i* :: *finite* ⇒ *'i pmf* **where**
transition-of-st-mat a i = *pmf-as-measure.pmf-of-st-vec* (*transition-vec-of-st-mat a i*)

lemma *st-vec-transition-vec-of-st-mat*[*simp*]:
st-vec (*transition-vec-of-st-mat A a*) $ *i* = *st-mat A* $ *i* $ *a*
⟨*proof*⟩

locale *transition-matrix* = *pmf-as-measure* +
fixes *A* :: *'i* :: *finite st-mat*
begin
sublocale *MC-syntax transition-of-st-mat A* ⟨*proof*⟩

lemma *measure-pmf-of-st-vec*[*simp*]: *measure-pmf* (*pmf-of-st-vec x*) = *measure-of-st-vec x*
⟨*proof*⟩

lemma *pmf-transition-of-st-mat*[*simp*]: *pmf* (*transition-of-st-mat A a*) *i* = *st-mat A* $ *i* $ *a*
⟨*proof*⟩

lemma *bind-is-matrix-vector-mult*: (*bind-pmf x* (*transition-of-st-mat A*)) =
pmf-as-measure.pmf-of-st-vec (*A* ∗*st st-vec-of-pmf x*)
⟨*proof*⟩

lemmas *stationary-distribution-alt-def* =
stationary-distribution-def[*unfolded bind-is-matrix-vector-mult*]

lemma *stationary-distribution-implies-pmf-of-st-vec*:
assumes *stationary-distribution N*
shows ∃ *x*. *N* = *pmf-of-st-vec x*
⟨*proof*⟩

lemma *stationary-distribution-pmf-of-st-vec*:
stationary-distribution (*pmf-of-st-vec x*) = (*A* ∗*st x* = *x*)
⟨*proof*⟩
end
end

5 Eigenspaces

Using results on Jordan-Normal forms, we prove that the geometric multiplicity of an eigenvalue (i.e., the dimension of the eigenspace) is bounded by the algebraic multiplicity of an eigenvalue (i.e., the multiplicity as root of the characteristic polynomial.). As a consequence we derive that any two eigenvectors of some eigenvalue with multiplicity 1 must be scalar multiples of each other.

theory *Eigenspace*
imports
Jordan-Normal-Form.Jordan-Normal-Form-Uniqueness
Perron-Frobenius.Perron-Frobenius-Aux
begin
hide-const (**open**) *Coset.order*

The dimension of every generalized eigenspace is bounded by the algebraic multiplicity of an eigenvalue. Hence, in particular the geometric multiplicity is smaller than the algebraic multiplicity.

lemma *dim-gen-eigenspace-order-char-poly*: **assumes** *jnf*: *jordan-nf A n-as*
shows *dim-gen-eigenspace A lam k* ≤ *order lam* (*char-poly A*)
⟨*proof*⟩

Every eigenvector is contained in the eigenspace.

lemma *eigenvector-mat-kernel-char-matrix*: **assumes** *A*: *A* ∈ *carrier-mat n n*
and *ev*: *eigenvector A v lam*
shows *v* ∈ *mat-kernel* (*char-matrix A lam*)
⟨*proof*⟩

If the algebraic multiplicity is one, then every two eigenvectors are scalar multiples of each other.

lemma *unique-eigenvector-jnf*: **assumes** *jnf*: *jordan-nf* (*A* :: ′*a* :: *field mat*) *n-as*
and *ord*: *order lam* (*char-poly A*) = *1*

and *ev*: *eigenvector A v lam eigenvector A w lam*
shows ∃ *a*. *v* = *a* $\cdot_v$ *w*
⟨*proof*⟩

Getting rid of the JNF-assumption for complex matrices.

lemma *unique-eigenvector-complex*: **assumes** *A*: *A* ∈ *carrier-mat n n*
and *ord*: *order lam* (*char-poly A* :: *complex poly*) = *1*
and *ev*: *eigenvector A v lam eigenvector A w lam*
shows ∃ *a*. *v* = *a* $\cdot_v$ *w*
⟨*proof*⟩

Convert the result to real matrices via homomorphisms.

lemma *unique-eigenvector-real*: **assumes** *A*: *A* ∈ *carrier-mat n n*
and *ord*: *order lam* (*char-poly A* :: *real poly*) = *1*
and *ev*: *eigenvector A v lam eigenvector A w lam*
shows ∃ *a*. *v* = *a* $\cdot_v$ *w*
⟨*proof*⟩

Finally, the statement converted to HMA-world.

lemma *unique-eigen-vector-real*: **assumes** *ord*: *order lam* (*charpoly A* :: *real poly*) = *1*
and *ev*: *eigen-vector A v lam eigen-vector A w lam*
shows ∃ *a*. *v* = *a* ∗*s* *w* ⟨*proof*⟩

end

6 Stochastic Matrices and the Perron–Frobenius Theorem

Since a stationary distribution corresponds to a non-negative real eigenvector of the stochastic matrix, we can apply the Perron–Frobenius theorem. In this way we easily derive that every stochastic matrix has a stationary distribution, and moreover that this distribution is unique, if the matrix is irreducible, i.e., if the graph of the matrix is strongly connected.

theory *Stochastic-Matrix-Perron-Frobenius*
imports
Perron-Frobenius.Perron-Frobenius-Irreducible
Stochastic-Matrix-Markov-Models
Eigenspace
begin

hide-const (**open**) *Coset.order*

lemma *pf-nonneg-mat-st-mat*: *pf-nonneg-mat* (*st-mat A*)
⟨*proof*⟩

lemma *stoch-non-neg-vec-norm1*: **assumes** *stoch-vec* $(v :: real \,\widehat{}\, 'n)$ *non-neg-vec* v
  shows *norm1* $v = 1$
  ⟨*proof*⟩

lemma *stationary-distribution-exists*: $\exists\ v.\ A *st\ v = v$
⟨*proof*⟩

lemma *stationary-distribution-unique*:
  assumes *fixed-mat.irreducible* (*st-mat* A)
  shows $\exists!\ v.\ A *st\ v = v$
⟨*proof*⟩

Let us now convert the stationary distribution results from matrices to Markov chains.

context *transition-matrix*
begin

lemma *stationary-distribution-exists*:
  $\exists\ x.$ *stationary-distribution* (*pmf-of-st-vec* x)
⟨*proof*⟩

lemma *stationary-distribution-unique*: **assumes** *fixed-mat.irreducible* (*st-mat* A)
  shows $\exists!\ N.$ *stationary-distribution* N
⟨*proof*⟩
end
end

References

[1] J. Divasón, O. Kunar, R. Thiemann, and A. Yamada. Perron-frobenius theorem for spectral radius analysis. *Archive of Formal Proofs*, May 2016. http://isa-afp.org/entries/Perron_Frobenius.html, Formal proof development.

[2] J. Hölzl and T. Nipkow. Markov models. *Archive of Formal Proofs*, Jan. 2012. http://isa-afp.org/entries/Markov_Models.html, Formal proof development.