Izvestiya NAN Armenii. Matematika, vol. 44, no. 2, 2009, pp. 59-66.

MARKOV CHAINS, GAUSS SOUPS, AND COMPROMISE DYNAMICS

U. KRAUSE

University of Bremen, Bremen, Germany

ABSTRACT. It must have been around 1966 when I first met Klaus Krickeberg at a conference in Greece. In my memory it was Konrad Jacobs who encouraged some young people from Erlangen to attend. I still remember when Klaus took us, M. Sieveking, J. Köhn and me, to a quite exciting tour by car on the Peloppenese. Later on, times were exciting too when I was around 1970 assistent to Klaus in Heidelberg. I also remember David G. Kendall at the Greece-conference, telling us about the Delphi method. About this I had forgotten until recently when writing this paper I noticed that my current work on opinion dynamics is related to the Delphi method.

Dedicated to the 80th birthday of Klaus Krickeberg

1. INTRODUCTION

Markov chains represent an important tool used in mathematics, including statistics and stochastics, as well as in many other fields, ranging from physics over biology to sociology. Mathematically, a **Markov chain** is a discrete-time dynamical system

$$x(t+1) = Ax(t) \quad \text{for} \quad t = 0, 1, 2, \dots \quad \text{and} \quad x(0) \in \triangle_n, \tag{1.1}$$

where $x(t)$ denotes the state vector at time t, $\triangle_n$ the simplex

$$\triangle_n = \left\{ y \in \mathbb{R}^n_+ \mid \sum_{i=1}^{n} y_i = 1 \right\}, \qquad \mathbb{R}^n_+ \quad \text{the first orthant in } \mathbb{R}^n$$

and A the **transition matrix** with $a_{ij} \geq 0$ for $1 \leq i, j \leq n$ and $\sum_{i=1}^{n} a_{ij} = 1$ for all j. The most relevant single result about Markov chains is the following **Ergodic Theorem for primitive Markov Chains** (see [14], Theorem 4.2, p. 91) or **Basic Limit Theorem for Markov Chains** (see [13], Theorem p. 230) which states for a **primitive** or regular matrix A, that is some power of A has all its elements (strictly) positive, that

$$\lim_{t \to \infty} A^t = B. \tag{1.2}$$

Thereby, all columns of matrix B are equal and given by the vector $q \in \triangle_n$ uniquely determined by $Aq = q$. It is common to prove this theorem by considering Markov chains as particular positive dynamical systems and applying the Perron-Frobenius Theorem on nonnegative matrices (see for example [13] and [14]). Instead of assuming A to be primitive it is sufficient to assume that a power of A is **scrambling**, where the latter means that any two rows have a (strictly) positive value in a joint column. Although it seems not to be widely known, this condition is sufficient and necessary for (1.2) to hold. Below the Basic Limit Theorem for Markov Chains will be featured as a special case within the general framework of compromise dynamics.

In Section 2 Markov chains will be viewed as the special linear case of a so called Gaußsoup. The question of a Basic Limit Theorem for Gaußsoups will be raised and some answers will be given by referring to the literature.

In Section 3 Markov chains and Gauss soups will be embedded into the general frame work of compromise dynamics. A Basic Limit Theorem will be presented which covers those for Markov chains and Gauss soups.

In Section 4 open problems will be addressed which concern the dependence of the limit state vector on the initial state vector. Whereas for Markov chains this dependence is extremely simple it can be extremely difficult for Gauss soups. In the case of $n = 2$, the limit state vector which is given by the arithmetic-geometric mean depends on initial conditions via an elliptic integral, as was already observed by Gauss (see [2], [6]).

2. FROM MARKOV CHAINS TO GAUSS SOUPS

For a Markov chain the entry a_{ij} of the transition matrix A is interpreted as the probability to change from state j to state i. The equation $x_i(t+1) = \sum_{j=1}^n a_{ij}x_j(t)$ according to (1.1)then gives the probability for the system to be in state i at period $t+1$ in dependence of the probabilities for the previous period. Consider for the transposed matrix A^T the system dual to (1.1)

$$y(t+1) = A^T y(t) \quad \text{for} \quad t = 0, 1, 2, \ldots \quad \text{and} \quad y(0) \in \mathbb{R}^n_+. \tag{2.1}$$

The equation $y_i(t+1) = \sum_{j=1}^n a_{ji}y_j(t)$ means because of $\sum_{j=1}^n a_{ji} = 1$ that the state $y_i(t+1)$ is given as a weighted arithmetic mean of the states for the previous period. Considering the Basic Limit Theorem, property (1.2) is equivalent to $\lim_{t\to\infty}(A^T)^t = B^T$. Thus, we may use (1.1) and think in terms of probabilities or we may use (2.1) and think in terms of averaging. In the latter case we obtain for an initial vector $y(0)$ that

$$\lim_{t\to\infty} y(t) = \lim_{t\to\infty} (A^T)^t y(0) = B^T y(0) = c \tag{2.2}$$

where all components of vector c are equal. This can be interpreted as follows. Imagine a round of experts $1, 2, \ldots, n$ who in discussing some issue revise opinions according to system (2.1). That is each expert forms his opinion by taking a weighted arithmetic mean of the opinions of all experts in the previous period. Equation (2.2) then says that under the assumption of the Basic Limit Theorem for Markov chains the round of experts will approach a consensus. (See [4] for this point of view. Related are the earlier debates about the various Delphi methods as well as the recent concepts of prediction markets.) This interpretation leads immediately to the following question: Why should experts average by weighted arithmetic means and not, say, by weighted geometric means? This leads to the following concepts of a **Gauss soup** defined as a discrete-time dynamical system for $1 \leq i \leq n, t = 0, 1, 2, \ldots$

$$x_i(t+1) = f_i(x_i(t), \ldots, x_n(t)) \text{ for } x_1(0) > 0, \ldots, x_n(0) > 0 \tag{2.3}$$

where for weights $a_{ij} \geq 0$ with $\sum_{j=1}^n a_{ij} = 1$ either $f_i(x_1, \ldots, x_n) = \sum_{j=1}^n a_{ij} x_j$ or $f_i(x_1, \ldots, x_n) = \prod_{j=1}^n x_j^{a_{ij}}$. An interesting question is under what conditions on the matrix $A = (a_{ij})$ a Basic Limit Theorem does hold, that is the system (2.3) does for any initial conditions converge to a consensus. Consider the following simple example for $n = 2$, studied already by Gauss in 1799 (see [2], [6]):

$$x_1(t+1) = \frac{1}{2}(x_1(t) + x_2(t)), x_2(t+1) = \sqrt{x_1(t) x_2(t)} \text{ with } x_1(0) > 0, x_2(0) > 0.$$

It is easy to see that the system converges to a consensus c but it is not so easy to determine the arithmetic-geometric mean c in terms of initial conditions. A Gauss soup can be interpreted by saying that all experts average by a weighted arithmetic mean but whereas some possess a utility function which is just the identity others possess the logarithm as utility function.

A system of type (2.3) has been considered in [5] and [7] for **means of order r**. Given real numbers $p_i \geq 0$ with $\sum_{i=1}^n p_i = 1$ the latter is defined as $\prod_{i=1}^n x_i^{p_i}$ for $r = 0$ and as $(\sum_{i=1}^n p_i x_i^r)^{\frac{1}{r}}$ for $r \in \mathbb{R}, r \neq 0$.

The following result is obtained in [5] and [7] by different reasoning; there is no reference to [7] in [5].

Proposition Let in (2.3) $f_i(x_1, \ldots, x_n)$ be a mean of order r_i for $1 \leq i \leq n$. If $p_i > 0$ holds for all i then

$$\lim_{t \to \infty} x_i(t) = c \quad \text{for all} \quad 1 \leq i \leq n$$

where c may depend on the initial conditions.

This proposition as well as a similar result for Gauss soups we will obtain in the next section as special cases within a much broader framework.

3. FROM GAUSS SOUPS TO COMPROMISE DYNAMICS

The (dual) Markov chain of equation (2.1) and, more general, a Gauss soup as defined by equation (2.3) has been interpreted as a process of opinion formation among experts where each expert takes others into account by using a weighted arithmetic or geometric mean. One may think also of experts using means of order r. All these means can be looked at just as different ways of making a compromise. Up to now opinions were represented by real numbers but we may allow for opinions having several dimensions. Let $S \subseteq \mathbb{R}^d$, for $d \geq 1$, denote the set of all possible opinions. What then could a compromise mean in this setting? An obvious idea would be to model a **compromise** of opinions $x^1, x^2, \ldots, x^n \in S$ as a convex combination $\alpha_1 x^1 + \alpha_2 x^2 + \ldots + \alpha_n x^n$ with $0 \leq \alpha_i \leq 1$ and $\sum_{i=1}^n \alpha_i = 1$. Thus, let S be a convex set and denote the set of all convex combinations by conv $\{x^1, \ldots, x^n\}$. A map $f : S^n \to S^n$ is called a **compromise map** on S for n experts if the component maps f_i of f satisfy the condition

$$\operatorname{conv}\{f_1(x), \ldots, f_n(x)\} \subseteq \operatorname{conv}\{x^1, \ldots, x^n\} \tag{3.1}$$

for all $x = (x^1, \ldots, x^n) \in S^n$.
If there holds already a consensus c, that is $x^1 = \ldots = x^n = c$, then (3.1) implies that $f_1(x) = \ldots = f_n(x) = c$. A compromise map is called **proper** if in all other cases, that is for $x^1, \ldots, x^n$ not all equal, the inclusion in (3.1) holds properly, $\subseteq$ but not $=$. The following result on compromise maps is obtained in [11].

Theorem For a proper and continuous compromise map f the compromise algorithm given by the recursion

$$x^i(t+1) = f_i(x^1(t), \ldots, x^n(t)) \quad \text{for} \quad 1 \leq i \leq n, t \in \{0, 1, 2, \ldots\} \tag{3.2}$$

converges always to a consensus. That is, there exists $c = c(x(0))$ such that

$$\lim_{t \to \infty} x^i(t) = c \text{ for all } 1 \leq i \leq n, x(0) = (x^1(0), \ldots, x^n(0)) \in S^n. \tag{3.3}$$

The theorem can be strengthened by the following

Addendum Instead of assuming in the Theorem the compromise map f to be proper it suffices to assume that some iterate f^k of f is proper.

Proof. Obviously, f^k is a compromise map, too. By the Theorem

$$\lim_{t \to \infty} f^{kt}(x) = C(x), \quad C(x) = (c(x), \ldots, c(x)), \quad x = x(0).$$

By continuity of f

$$\lim_{t \to \infty} f^{kt+1}(x) = f(\lim_{t \to \infty} f^{kt}(x)) = f(C(x)).$$

From (3.1) it follows that

$$f(C(x)) = (f_1(C(x)), \dots, f_n(C(x))) = (c(x), \dots, c(x)) = C(x).$$

By iteration

$$\begin{aligned} \lim_{t\to\infty} f^{kt+i}(x) &= C(x) \quad \text{for} \quad 1 \le i \le k-1 \qquad \text{and, hence,} \\ \lim_{s\to\infty} f^{s}(x) &= C(x). \end{aligned}$$

□

For the special case of one dimension, $d = 1$, we obtain the following result.

Corollary 1 Let S be an interval in $\mathbb{R}$ and $f : S^n \to S^n$ be a continuous map such that for $1 \le i \le n$ and $x \in S^n$

$$\min_{1\le j\le n} x^j \le f_i(x) \le \max_{1\le j\le n} x^j. \tag{3.4}$$

Suppose, for some iterate f^k and given $x = (x^1, \dots, x^n)$ with not all components equal at least one of the two inequalities in (3.4), with $f_i(x)$ replaced by $f_i^k(x)$, holds strictly for all $1 \le i \le n$. Then the compromise algorithm for f converges always to a consensus.

Proof. With $a(x) = \min\limits_{1\le j\le n} x^j$ and $b(x) = \max\limits_{1\le j\le n} x^j$ we have that $\mathrm{conv}\{x^1, \dots, x^n\}$ is the closed interval $[a(x), b(x)]$. Thus, (3.4) implies that f is a compromise map on S. Obviously, f^k is a compromise map, too. By assumption for x with not all components equal $f_i^k(x) \in]a(x), b(x)]$ for all i or $f_i^k(x) \in [a(x), b(x)[$ for all i. Therefore, f^k is proper and the Addendum applies. □

One verifies easily that all the means mentioned satisfy inequalities (3.4) (for $S = \mathbb{R}_+$). Indeed, (3.4) describes what a mean "means". Such a mapping is called a *generalized mean* (or abstract mean) in [2].

There it is shown that for a mapping with components satisfying (3.4) the iterates converge to consensus, provided both inequalities in (3.4) hold strictly for x with not all components equal. The latter conditions is more demanding than the one in Corollary 1 as can be seen from the following example $f : \mathbb{R}_+^2 \to \mathbb{R}_+^2, f_1(x_1, x_2) = \min\{x_1, x_2\}, f_2(x_1, x_2) = \frac{1}{2}(x_1 + x_2)$.

Corollary 1 applies to the following joint extension of Gauss soups and mappings with components given by means of order r. Let $f : \mathbb{R}_+^n \to \mathbb{R}_+^n (\mathbb{R}_+$ without 0) be a mapping where for each $1 \le i \le n$ the component mapping $f_i(x_1, \dots, x_n)$ is given either by

$$\left(\sum_{j=1}^{n} a_{ij} x_j^{r_i}\right)^{\frac{1}{r_i}} \qquad \text{or by} \qquad \prod_{j=1}^{n} x_j^{a_{ij}}. \tag{3.5}$$

Thereby, for all $i, j, r_i \neq 0, 0 \leq a_{ij}, \sum_{j=1}^n a_{ij} = 1$. The matrix $A = (a_{ij})$ is called **scrambled** if for any two rows i and j there exists a column k, such that $a_{ik} > 0$ and $a_{jk} > 0$ (see [14]).

Corollary 2 The compromise algorithm for the map f defined above by (3.5) converges always to consensus provided the matrix $A = (a_{ij})$ is scrambled.

Proof. Obviously, the inequalities (3.4) are satisfied for each f_i. To obtain the result from Corollary 1 we show that the equations $f_i(x) = \min_\ell x^\ell$ and $f_j(x) = \max_\ell x^\ell$ for any two indices i and j imply that all components of x must be equal. We consider the case where

$$f_i(x) = \left(\sum_{h=1}^n a_{ih} x_h^{r_i}\right)^{\frac{1}{r_i}} \qquad \text{and} \qquad f_j(x) = \prod_{h=1}^n x_h^{a_{jh}} \ ;$$

the other cases can be treated the same way. It follows

$$\sum_{h=1}^n a_{ih} x_h^{r_i} = \left(\min_\ell x^\ell\right)^{r_i} \qquad \text{and} \qquad \prod_{h=1}^n x_h^{a_{jh}} = \max_\ell x^\ell$$

and, hence

$$\begin{aligned} \sum_{h=1}^n a_{ih}\left(x_h^{r_i} - (\min_\ell x^\ell)^{r_i}\right) &= 0 \qquad\qquad \text{and} \\ \prod_{h=1}^n \left(\frac{x_h}{\max_\ell} x^\ell\right)^{a_{jh}} &= 1\ . \end{aligned}$$

By assumption there exists k such that $a_{ik} > 0$ and $a_{jk} > 0$. The former implies that $x_k^{r_i} - \left(\min_\ell x^\ell\right)^{r_i} = 0$ and the latter that $\frac{x_k}{\max_\ell \ x^\ell} = 1$. Thus $\min_\ell \ x^\ell = x_k = \max_\ell \ x^\ell$ which means that all x_ℓ must be equal. □

The Proposition stated in Section 2 is a special case of Corollary 2 by observing that the matrix A having all its rows equal to the vector $(p_1, \ldots, p_n)$ is strictly positive and, hence, scrambled in a trivial way.

Since a Markov chain is a special case of the system given by (3.5), Corollary 2 applies to Markov chains, too. But in this case we can obtain a sharper result from Corollary 1.

Corollary 3 The Basic Limit Theorem for a Markov chain given by the matrix A holds provided some power of A is scrambled.

Proof. The map defined by $f_i(x) = \sum_{j=1}^n a_{ij} x_j$ satisfies (3.4) and f_i^k satisfies the assumptions made in Corollary 1 by the same argument as in the proof of Corollary 2 when applied to the power A^k. □

Thus, looking from Corollary 3 back at the Theorem and its Addendum in Section 3 the latter may be viewed as an extension of the Basic Limit Theorem for Markov Chains to the more general compromise maps.

There are many more consequences of the Theorem on compromise dynamics. It can be applied to a nonlinear model of opinion dynamics developed in [9] and [10] and also to a model of collective dynamics in [12] where experts are equipped with utility functions.

4. OPEN PROBLEMS

A quite natural question with respect to the compromise algorithm of the Theorem is the question of how the consensus $c(x(0))$ does depend on the initial conditions $x(0)$. For the case of a Markov chain where the compromise map is given by $f(y) = A^T y$ this question is easily answered. For A scrambling from (2.2) we have that

$$\lim_{t\to\infty} y(t) = \lim_{t\to\infty} (A^T)^t y(0) = B^T y(0).$$

Since all columns of B are given by the vector $q \in \triangle_n$ given uniquely by $Aq = q$ it follows for the consensus that

$$c(y(0)) = \sum_{j=1}^{n} q_j y(0)_j. \tag{4.1}$$

In particular, the consensus is a linear function of the initial conditions.

What about nonlinear compromise maps, in particular Gauss soups? Consider the simple case of the Gauss soup given by $f_1(x_1, x_2) = \frac{x_1+x_2}{2}$, $f_2(x_1, x_2) = \sqrt{x_1 x_2}$. Gauss proved that

$$c(x_1(0), x_2(0)) = \frac{\pi}{2}\left[\int_0^{\frac{\pi}{2}} \frac{d\varphi}{\sqrt{x_1(0)^2\cos^2\varphi + x_2(0)^2\sin^2\varphi}}\right]^{-1}.$$

Surprisingly, the dependency of the consensus on initial conditions is quite involved. What about a Gauss soup with other weights in case of $n = 2$, that is

$$f_1(x_1, x_2) = \alpha x_1 + (1-\alpha)x_2, \quad f_2(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha} \qquad ?$$

I do not know of any explicit formula for $c(x_1(0), x_2(0))$.

What about $n = 3$, Gauss soups with three experts? I simply do not know. Gauss himself and many mathematicians following him explored a lot of special cases. For example for $n = 2$ the proper compromise map given by

$$f_1(x_1, x_2) = \frac{x_1 + x_2}{2}, \quad f_2(x_1, x_2) = \sqrt{\frac{x_1 + x_2}{2}\, x_2}$$

or for $n = 3$ the proper compromise map given by

$$f_1(x_1, x_2, x_3) = \frac{x_1 + x_2 + x_3}{2},$$

$$f_2(x_1, x_2, x_3) = \sqrt{\frac{x_1x_2 + x_1x_3 + x_2x_3}{3}} \tag{4.2}$$

$$f_3(x_1, x_2, x_3) = \sqrt[3]{x_1x_2x_3}.$$

For $n = 4$ Borchardt extended Gauss' formula (4.2) for the proper compromise map with components given by

$$\frac{1}{4}(x_1+x_2+x_3+x_4), \frac{1}{2}(\sqrt{x_1x_2}+\sqrt{x_3x_4}), \frac{1}{2}(\sqrt{x_1x_3}+\sqrt{x_2x_4}), \frac{1}{2}(\sqrt{x_1x_4}+\sqrt{x_2x_3}). \tag{4.3}$$

(For these and many more examples see [1], [2], [3].)

Open problems abundant also for compromise maps covered by Corollary 2, for example the following variation of the simplest proper Gauss soup

$$f_1(x_1, x_2) = \left(\frac{x_1^r + x_2^r}{2}\right)^{\frac{1}{r}}, \quad r \neq 0, \quad f_2(x_1, x_2) = \sqrt{x_1x_2}. \tag{4.4}$$

Except for $r = 1$, I do not know of any explicit formula for $c(x_1(0), x_2(0))$.

References

[1] C. W. Borchardt, "Theorie des arithmetisch-geometrischen Mittels aus vier Elementen", Math. Abh. Berliner Akademie der Wissenschaften, 33-96 (1878).

[2] J. M. Borwein and P. B. Borwein, *Pi and the AGM. A Study in Analytic Number Theory and Computational Complexity.* (John Wiley & Sons, New York, 1987).

[3] B.C. Carlson, "Algorithms involving arithmetic and geometric means", Amer. Math. Monthly, **78**, 496-505 (1971).

[4] M. H. De Groot, "Reaching a consensus", J. Amer. Statist. Assoc., **69**, 118-121 (1974).

[5] C. J. Everett and N. Metropolis, "A generalization of the Gauss limit for iterated means", Advances in Mathematics, **7**, 297-300 (1971).

[6] C. F. Gauss, "Werke", **3**, 361-404 (Göttingen 1866/1933).

[7] W. Gustin, "Gaussian means", Amer. Math. Monthly, **54**, 332-335 (1947).

[8] R. Hegselmann and U. Krause, "Opinion dynamics and bounded confidence: models, analysis and simulation", Journal of Artificial Societies and Social Simulation, **5**, (3) (2002). http://jasss.soc.surrey.ac.uk/5/3/2.html.

[9] R. Hegselmann and U. Krause, "Opinion dynamics driven by various ways of averaging", Computational Economics, **25**, 381-401 (2005).

[10] U. Krause, "Arithmetic-geometric discrete systems", J. Difference Equat. Appl., **12**, 229-231 (2006).

[11] U. Krause, "Compromise, consensus and the iteration of means", Elemente der Mathematik, forthcoming.

[12] U. Krause, "Collective dynamics of Faustian agents", to appear.

[13] D. G. Luenberger, *Introduction to Dynamic Systems.* (Wiley & Sons, New York, 1979).

[14] E. Seneta, *Non-negative Matrixes and Markos Chains.* (2nd edition. Springer, New York 1980).

January 20, 2009