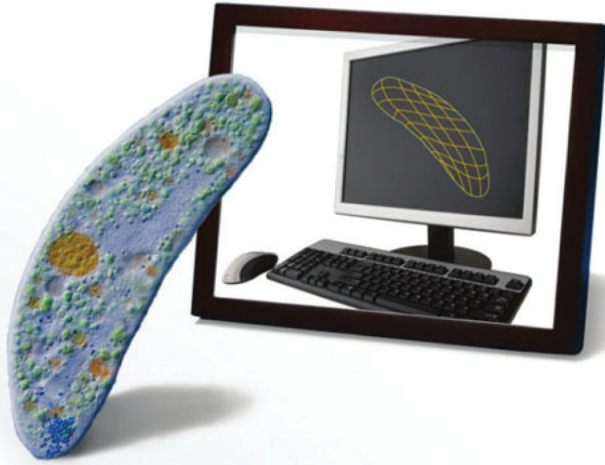


# 수리생물학 (Mathematical Biology)

김재경 KAIST 수리과학과

원자(Atom)로 구성된 물리적 세계를 비트(bit)로 구성된 가상 세계에 구현하는 4차 산업혁명이 급속하게 일어나고 있다. 생명과학 분야 역시 생명 체계를 디지털 공간에 가상으로 구현하여 복잡한 데이터를 분석하고 가상 실험을 하는 혁명이 진행 중이다(그림 1). 이를 위한 수학 이론을 개발하고 응용하는 분야를 수리생물학(Mathematical Biology)이라고 한다.

국내엔 아직 낯선 분야이지만 최근 들어 미국 수학 박사 6명 중 1명, 통계학 박사의 절반은 생물학 관련 연구로 학위를 받을 정도로 급속하게 성장하고 있는 분야이다. 지난 5월 미국 연구 재단(National Science Foundation)과 사이먼 재단(Simon Foundation)은 공동으로 매년 450억 원을 지원하는 수리생물학 연구소를 미국 동부·중부·서부·남부 네 곳에 동시에 설립하며 생명과학의 혁명을 위한 속도를 높이고 있다.

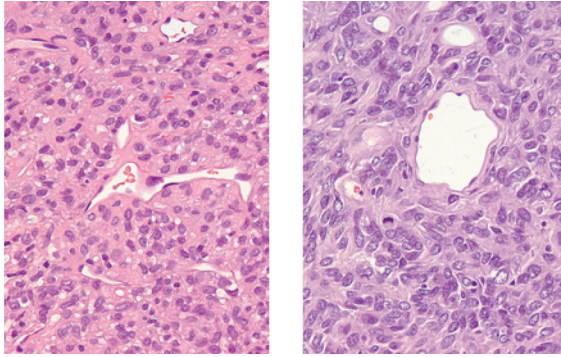


〈그림 1〉 생명 체계를 가상 세계에 구현하여 복잡한 데이터를 분석하고 가상 실험을 진행하는 생명과학의 혁명이 진행 중이다. 박테리아의 가상 모델 구현에 성공한 Markus Covert Lab(스탠퍼드대학교)에서 그림을 얻음.

## 데이터 분석

우리의 몸은 우주에 존재하는 모든 별보다 훨씬 많은 수인 약 100조 개 세포로 이루어졌고, 각 세포는 약 100조 개 분자들로 구성되어 있다. 이러한 복잡한 시스템을 컴퓨터의 도움 없이 인간의 직관만으로 이해하는 것은 불가능하다.

환자의 종양 세포 사진을 보고 안전한 양성 종양과 암이 되어 전이를 일으킬 악성 종양을 구별하는 것은 중요한 문제지만, 육안으로 판단하기는 쉽지 않다(그림 2). 이 문제에 접근하는 가장 단순한 방법은 이제까지 관찰된 수많은 종양 사진들에서 양성 종양과 악성 종양의 특징들을 각각 정리하여 외워두었다가 새로운 환자의 종양 사진이 어느 것과 더 비슷한지 판단하는 것일 터이다. 현재 세포의 거침 정도, 면적, 도드라짐 정도 등 30여 개 특징이 판단에 사용된다.



〈그림 2〉양성과 악성 종양의 사진. 육안으로 판별하기 어렵다.(DeVito et al, *PLoS One*, 2015)

행렬(Matrix)을 이용하면 이러한 판단을 자동화할 수 있다. 우선 기존에 알려진 각 종양의 다양한 특징을 정량화하여 행렬을 만든다(그림 3). 그리고 새로운 종양 사진을 정량화한 후 행렬의 어느 행과 가장 비슷한지 찾으면 새로운 종양의 종류를 판단할 수 있다.

그런데 비슷함을 어떻게 판단해야 할까? 가장 쉽게 생각할 수 있는 방법은 두 종양의 30여 가지 특징을 나타내는 값들의 차를 구한 후 그 절댓값의 합이 작을수록 비슷하다고 하는 것이다. 이 방식은 양성과 악성 종양을 판단하는 데 30여 가지 특징이 동등하게 중요하다고 전제한다. 만약 종양의

	Texture	Area	Compactness	Dimension	Radius	Concavity .....
Tumor 1	1	4	...			
Tumor 2	3.1	0.8	...			
.	.	.	.			
Tumor 1000	0.5	0.25	...			
.	.	.	.			
.	.	.	.			

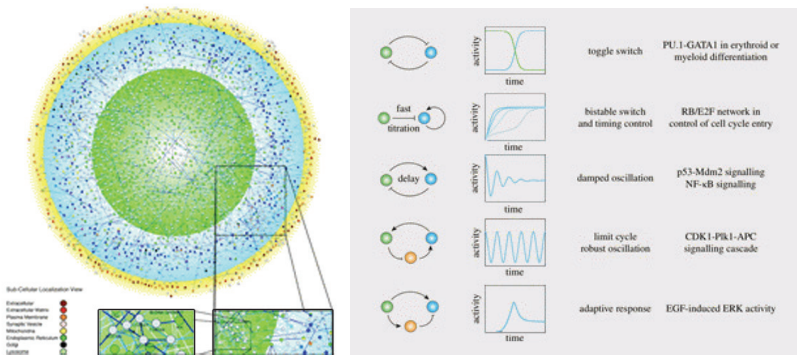
〈그림 3〉 종양 사진을 다양한 특징으로 정량화하여 행렬을 만들 수 있다. 데이터의 행렬화는 모든 데이터 분석의 기본이다.

크기가 양성과 악성의 판단에 그리 중요하지 않은 요소라면 이러한 방식은 오판을 일으킬 가능성이 높다. 그렇다면 양성과 악성 종양 판단에 중요한 요소들이 무엇인지 어떻게 찾을 수 있을까? 행렬 분해(Matrix Factorization)나 주성분 분석(Principal Component Analysis)과 같은 알고리즘을 행렬에 적용하면 양성과 악성의 판별에 중요한 요소들을 찾아낼 수 있다. 이러한 알고리즘을 실제로 위스콘신 유방암 진단 데이터(Wisconsin Breast Cancer Diagnostic Data Set)에 적용하면 양성과 악성을 판단할 수 있는 핵심 지표가 4~5개임을 알 수 있다. 나아가 행렬에 알파고(Alpha Go) 개발에 사용된 딥러닝(Deep Learning) 알고리즘을 적용하면 32가지 특징에 적절한 가중치를 부여해서 정교한 양성/악성 판단 알고리즘을 만들어낼 수 있다.

또 다른 예로 소아들의 자폐증(Autism Spectrum Disorder) 진단을 들 수 있다. 이 진단의 어려운 점 중 하나는 소아들에게서 빈번하게 발생하는 ADHD(Attention Deficit Hyperactivity Disorder)와 비슷한 증상이 많다는 것이다. 두 질병을 구별하기 위해 아이의 부모가 자녀의 행동에 대하여 답하는 SRS(Social Responsiveness Scale)가 널리 사용되어 왔다. 하지만 SRS는 65개 질문으로 구성되어 그 답을 분석하여 진단하는 데 시간이 많이 걸렸다. 최근 미국 스텐퍼드대학교 월(D. P. Wall) 교수 연구팀은 환자 3000명의 SRS 답변 결과(즉  $3000 \times 65$  행렬)를 다양한 기계 학습(Machine Learning) 알고리즘을 이용하여 분석하였다(M. Duda et al, *Translational Psychiatry*, 2016). 그 결과 65개 질문 중 단 5개 질문에 대한 답을 알면 자폐증과 ADHD를 구별할 수 있다는 것을 밝혀 자폐증 진단의 효율성과 정확성이 획기적으로 높아졌다.

## 가상 실험

노벨 물리학상 수상자인 에르빈 슈뢰딩거(Erwin Schrödinger)는 1944년에 발간된 저서 『생명이란 무엇인가(What is life?)』에서 죽음을 변화하지 않는 평형상태(Equilibrium)로, 생명은 끊임없이 변화하는 역동적인(Dynamic) 상태로 정의하였다.



〈그림 4〉(좌) 한 개 세포도 약 100조 개 분자(유전자, 단백질 등)로 이루어진다.(Giot et al, *Science*, 2003) (우) 이들은 생화학 반응을 통해 생명 유지에 필수적인 기능을 하는 다양한 역학(Dynamics)을 만들어낸다.(Dong and Liu, *Open Biology*, 2017)

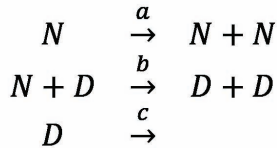
실제로 지난 수십 년간 진행된 분자생물학 혁명을 통해 세포 속 수많은 분자가 생화학 반응을 통해 끊임없이 변화한다는 것이 밝혀졌다(그림4). 세포분열 조절 분자들은 세포가 안정적으로 분열할 수 있는 조건에서만 작동하는 스위치 역학(Dynamics)을 통해서 돌연변이와 암을 예방한다. DNA 손상이 발생하면 암 억제 분자인 p53 단백질의 양이 주기적으로 증감을 반복하며 DNA 손상을 치료한다. Period 분자들은 24시간 주기로 증감을 반복하며 생명 시스템의 시간 장치 역할을 한다.<sup>[1]</sup> 온도 변화와 같은

[1] Period 유전자를 초파리에서 찾고 Period 단백질의 농도가 생체리듬에 맞춰 24시간마다 한 번씩 주기적으로 변화하는 것을 밝힌 공로로 제프리 C. 홀, 마이클 로스배시, 마이클 W. 영 교수가 2017년 노벨 생리학상을 받았다.

환경 변화의 스트레스를 감지하는 분자들은 스트레스가 발생할 때 적응(Adaptation) 반응 역학(Dynamics)을 만들어냄으로써 새로운 환경에 적응할 수 있도록 한다.

이러한 다양한 역학(Dynamics)이 유전적 변이 등의 이유로 멈추면 질병이 발생하고 심각한 경우 죽음에 이른다. 따라서 치료법 개발을 위해 어떻게 분자들이 다양한 역학(Dynamics)을 만들어내는지 이해하는 것이 중요하다. 하지만 분자들의 복잡한 상호작용이 어떻게 작동하는지 이해하는 것은 굉장히 어려운 문제이다. 이렇게 끊임없이 변화하는 생명 시스템을 이해하고 예측하기 위한 수학 이론 중 하나가 미분방정식(Differential Equations)이다.

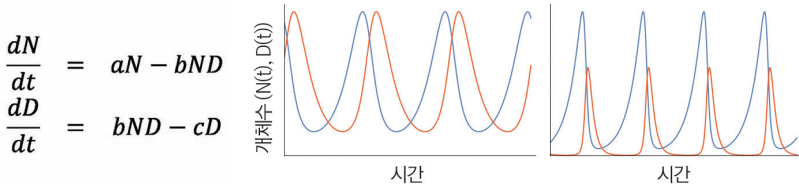
예를 들어 우리 몸속의 정상 세포(N)가 속도 a로 세포 분열하여 개체수가 두 배씩 늘어난다고 하자(그림 5). 그런데 몸속에 침투한 병원 세포(D)가 속도 b로 건강한 세포와 만나 건강한 세포를 감염시키고 병원 세포로 변형시킨다. 그리고 병원 세포는 속도 c로 소멸한다.



〈그림 5〉 정상 세포(N)와 병원 세포(D)의 상호작용을 나타내는 반응식

과연 별다른 치료를 하지 않아도 시간이 지나면 병원 세포가 소멸하여 정상 세포만 남을까? 아니면 정상 세포가 모두 감염되어 치료가 필요할까? 치료해야 한다면 정상 세포의 감염 속도(b)를 줄이는 약과 병원 세포의 소멸 속도(c)를 빠르게 하는 약 중 어느 것이 더 효과적일까?

이러한 질문에 답하기 위해서는 건강한 세포와 병원 세포의 개체 수가



〈그림 6〉(좌) 정상 세포(N)와 병균 세포(D) 간의 상호작용을 묘사하는 미분방정식. (중) 미분방정식의 해. 정상 세포(파란색)와 병균 세포(빨간색)는 시간이 지남에 따라 주기적으로 개체수가 증감을 반복한다. (우) 병균 세포의 사망률(c)을 높이는 약을 투여해도 여전히 증감을 반복한다는 것을 미분방정식을 통해 예측할 수 있다.

시간이 지남에 따라 어떻게 변화할지 예측해야 하는데 직관만으로는 쉽지가 않다. 하지만 정상 세포와 병균 세포 사이의 상호작용 과정을 묘사하는 미분방정식을 유도한 후 풀면 시간이 지남에 따라 정상 세포와 병균 세포의 개체 수가 증감을 주기적으로 반복한다는 것을 알 수 있다(그림 6).

따라서 병균 세포를 제거하기 위해서는 적절한 치료법이 필요한데 미분방정식의 매개변수를 조절하여 새로운 해를 구함으로써 약의 효과를 예측하는 가상 실험을 할 수 있다. 즉 미분방정식에서 병균 세포의 소멸 속도 c를 크게 만든 후 해를 구해보면 주기와 진폭만 변화할 뿐 여전히 증감을 반복한다는 것을 알 수 있다(그림 6). 병균 세포가 거의 사라진 것처럼 보였다고 다시 증가하기에 이러한 치료법을 사용하면 곧 재발하리라는 것을 예측할 수 있다. 나아가 미분방정식의 분석을 통해 유일하게 병균을 없애는 방법은 감염되는 속도 b를 0으로 만드는 방법밖에 없다는 것을 알 수 있다. 이러한 미분방정식을 이용한 가상 실험으로 병균을 빨리 죽이는 약이 아니라 감염을 막는 약을 개발해야 한다는 결론을 내릴 수 있다.

마찬가지로 훨씬 더 복잡한 분자들의 상호작용 역시 미분방정식을 이용해서 가상 실험을 할 수 있다. 필자가 글로벌 제약회사 화이자(Pfizer)와 공동으로 진행하는 연구에서는 생체 시계(Circadian clock)를 구성하는 분자들의 상호작용을 300여 개 미분방정식을 통해서 묘사하고, 이를 이

용해 생체 시계 조절 신약의 효과를 가상으로 실험하고 있다(Kim et al, *CPT:PSP*, 2013). 이러한 가상 실험을 통한 약의 효과 예측은 신약 개발 비용을 혁신적으로 줄이고 성공률을 높일 것으로 기대된다.

## 맺음말

수리생물학에는 거의 모든 분야의 수학(대수학, 위상수학 등)이 쓰이지만 본 글에서는 고등학교 때 배운 ‘행렬’과 ‘미적분학’과 관련된 내용을 주로 소개하였다. 고교 시절 행렬과 미적분학을 배울 때는 이렇게 쓸모가 많을 것이라곤 상상하지 못했다. 다행히 4차 산업혁명의 흐름에서 발생하는 중요한 문제들이 학생들에게 이러한 수학을 배워야 하는 좋은 동기가 될 것으로 기대한다. 미분으로 복잡한 체계를 묘사할 수 있고, 적분을 통해 미분으로 표현된 방정식의 해를 구함으로써 복잡한 시스템의 미래를 예측할 수 있다. 따라서 미적분은 실제 세계가 어떻게 변화할지 디지털 세계에서 예측하기 위한 핵심 도구가 될 것이다. 행렬은 빅데이터 분석의 기본이기에 빅데이터의 시대를 살아갈 현재 세대에게 필수적인 교양이 될 것이다, 하지만 누구나 배울 수 있던 행렬이 몇 년 전부터 고교 수학 교과서에서 사라져 영재고와 같이 특수한 학교에 다니는 소수 학생만 배울 수 있게 되어 안타깝다. 복잡한 행렬을 분석하는 행렬 분해나 주성분 분석과 같은 알고리즘을 이해하기 위한 필수적인 개념인 기하와 벡터 역시 고교 과정에서 축소해야 한다는 논의가 있다고 하니 걱정이 된다.

필자의 연구실은 수학자들로 구성되어 있지만 의학·약학·생명과학 연구실들과 협력 연구를 진행하며 전공을 초월한 문제들을 매일 마주한다. 한 분야의 이론과 도구만으로 해결할 수 없는 퍼즐을 협력 연구자들과 한 조각 한 조각 함께 채워가며 학제 간 연구의 중요성을 매일 경험한다. 올 여름 KAIST 수리생물학 연구실의 첫 졸업생은 미국 미시간대학교(Uni-



versity of Michigan, Ann Arbor)의 수학과가 아닌 의과대학 연구원이 되었고 UNIST 수리생물학 연구실의 첫 졸업생은 학위 후 1년 만에 홋카이도 의과대학(Hokkaido Medical School) 조교수로 임용되었다는 소식이 들려왔다. 내년 봄에는 국내에서도 이러한 소식들이 전해지기를, 수학과와 생명과학자, 의학자가 대화를 나누는 것이 어색하지 않은 날이 곧 오기를 소망해본다.

#### [ 참고문헌 ]

- DeVito, Nicholas, et al., "Clinical characteristics and outcomes for solitary fibrous tumor (SFT): a single center experience", *PLoS One*, 10, 10(2015): e0140362.
- Duda, M., et al., "Use of machine learning for behavioral distinction of autism and ADHD", *Translational psychiatry*, 6, 2(2017): e732.
- Giot, Loic, et al., "A protein interaction map of *Drosophila melanogaster*", *Science*, 2003.
- Dong, Peng, and Zhe Liu, "Shaping development by stochasticity and dynamics in gene regulation", *Open biology*, 7, 5(2017): 170030.
- Kim, Jae Kyoung, et al., "Modeling and validating chronic pharmacological manipulation of circadian rhythms" *CPT: pharmacometrics & systems pharmacology* 2, 7(2013): pp. 1~11.

#### 김재경

KAIST 수리과학과 부교수. 서울대학교 수학교육과를 졸업하고 미국 미시간대학교에서 수학 박사 학위를 받았다. 박사 학위 후 오하이오주립대학 수리생물학 연구센터에서 박사 후 연구원 과정을 마친 뒤 2015년 KAIST 수리과학과에 조교수로 부임하였다. 수학을 이용하여 생명과학과 의학 분야의 다양한 문제를 연구하여 《사이언스》, 《몰라클러 셀(Molecular Cell)》, 《미국 학술원 학술지》에 우수 논문들을 게재하였다. 대한수학회 상산 젊은 수학자상(2015), 한국을 빛낸 젊은 과학자 30인(2016), 한국산업응용수학회 젊은 연구자상(2017), Human Frontier Science Program Young Investigator Award(2017)를 수상하였고 2017 International Conference on Systems Biology 기조 강연자로 선정되었다.

원고의 무단 전재와 복제를 금하며,  
원고를 임의로 편집하거나 배포할 수 없습니다.

## 지식의지평 25

知平 2018

발행인 | 대우재단

편집인 | 김광억

편집위원 | 김병연, 김승환, 김혜숙, 노정혜, 이태수, 주경철, 한준, 김용준, 김두철

편집주간 | 박은진

발행일 | 2018년 11월 22일

☒ 『지식의 지평』은 대우재단이 국내 학문의 발전과 소통을 위해 발간합니다.

☒ 게재된 내용은 대우재단 및 편집진의 의견과 다를 수 있습니다.