

최적화 알고리즘과 투자공학

서울대학교 컴퓨터공학부
최적화/금융공학 연구실

문 병 로

목 차

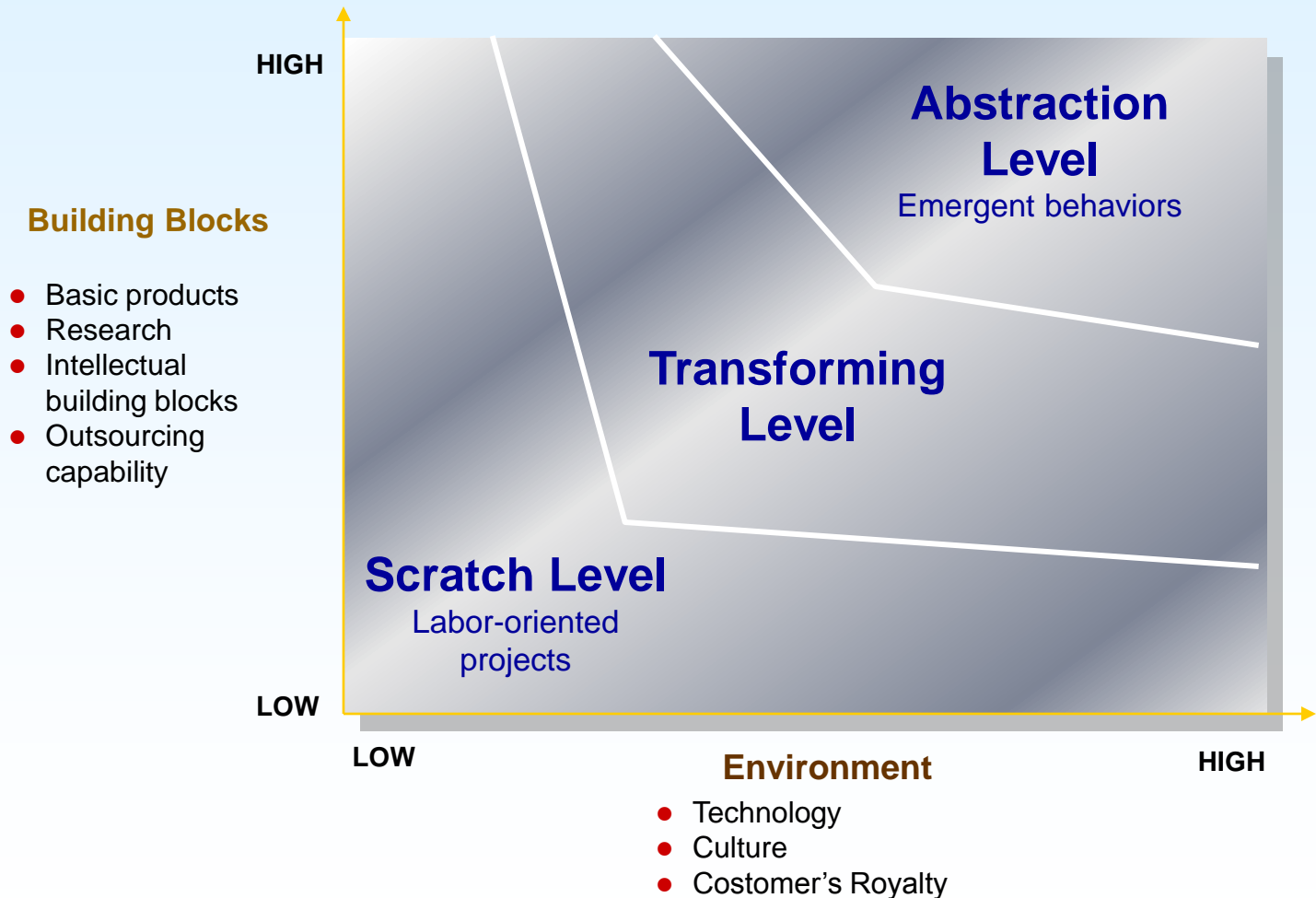
1. 도입, 문제공간
2. 최적화
3. 주식투자와 최적화

1. 도입

Motivations for Optimization

- **기능의 시대에서 효율성의 시대로!**
- 효율 지향 솔루션 시장의 확대
 - 금융
 - 자원 관리
 - 제조
 - CRM (고객관계관리)
 - SCM (공급망 관리)
 - 광고
 - 검색
 - ...
- 수요는 급증, 방법론은 정체
- 기법상의 심층적 변혁 필요

Company Intelligence Level



문제 공간 탐색

- Problem space

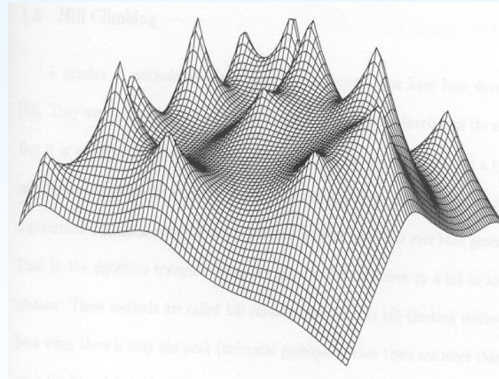
2-D



$$f(x) = \dots$$



3-D



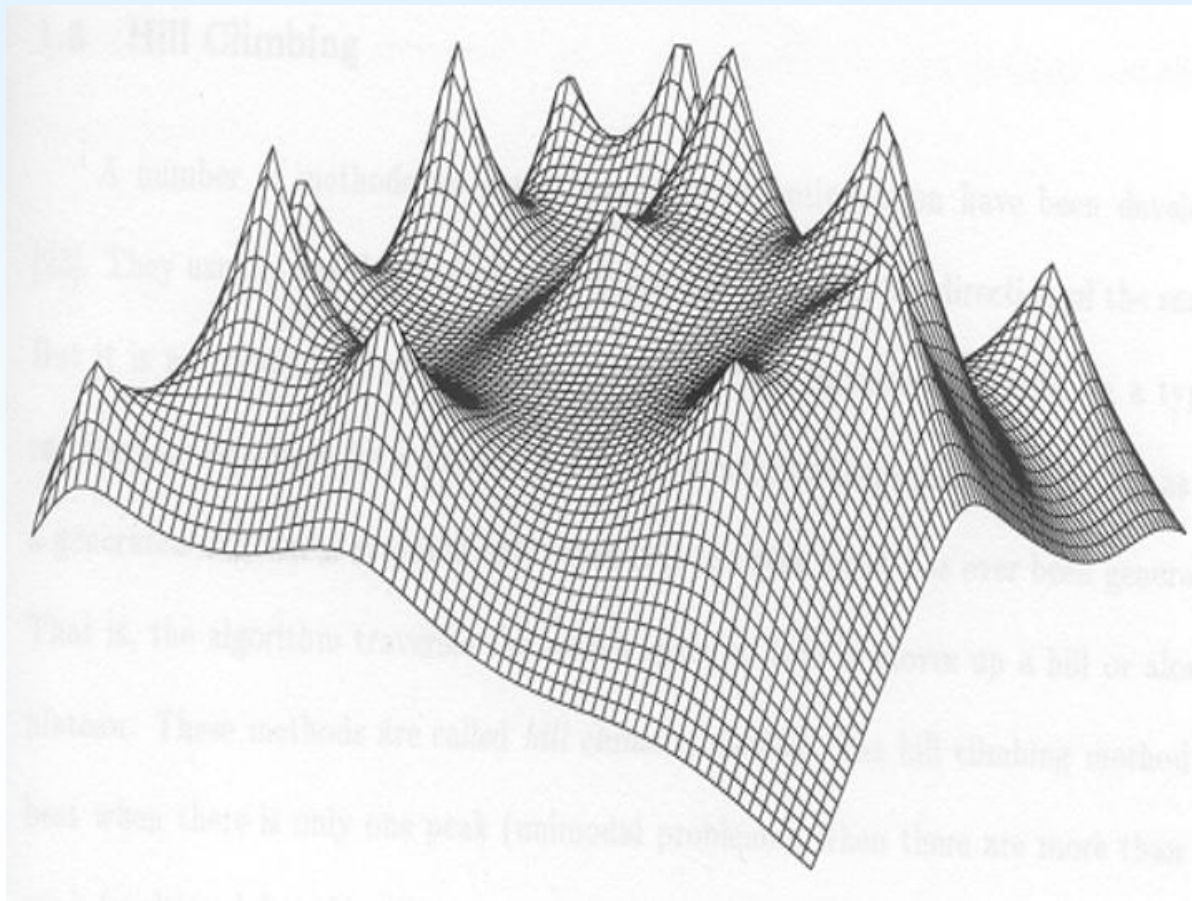
$$f(x, y) = \dots$$



N-D

?

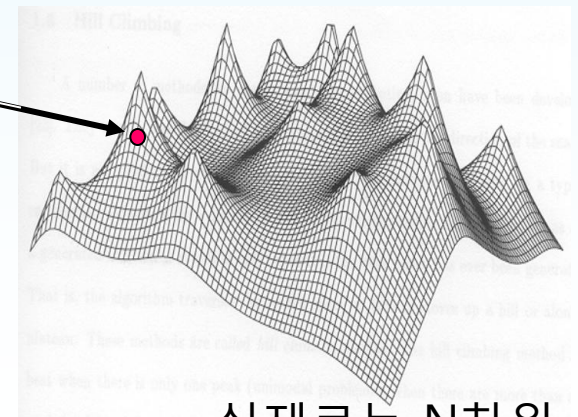
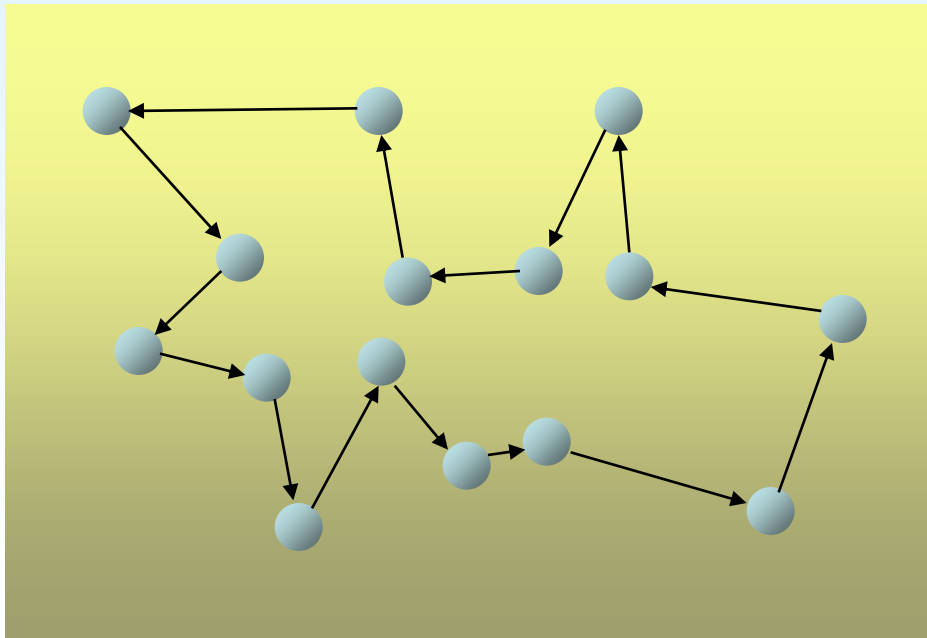
$$f(x_1, x_2, \dots, x_{N-1}) = \dots$$



- 각 격자점 하나는 하나씩의 솔루션
- 최적화
 - 문제공간에서 가장 높은 봉우리를 찾는 것
- 대부분 문제공간은 N 차원이다.

Traveling Salesman Problem (TSP)

- N 개의 도시가 주어지고,
모든 도시를 방문하고 돌아오는 최단경로를 찾는 문제
- 컴퓨터과학의 대표적 난제, NP-Hard



실제로는 N차원

문제공간의 방대함

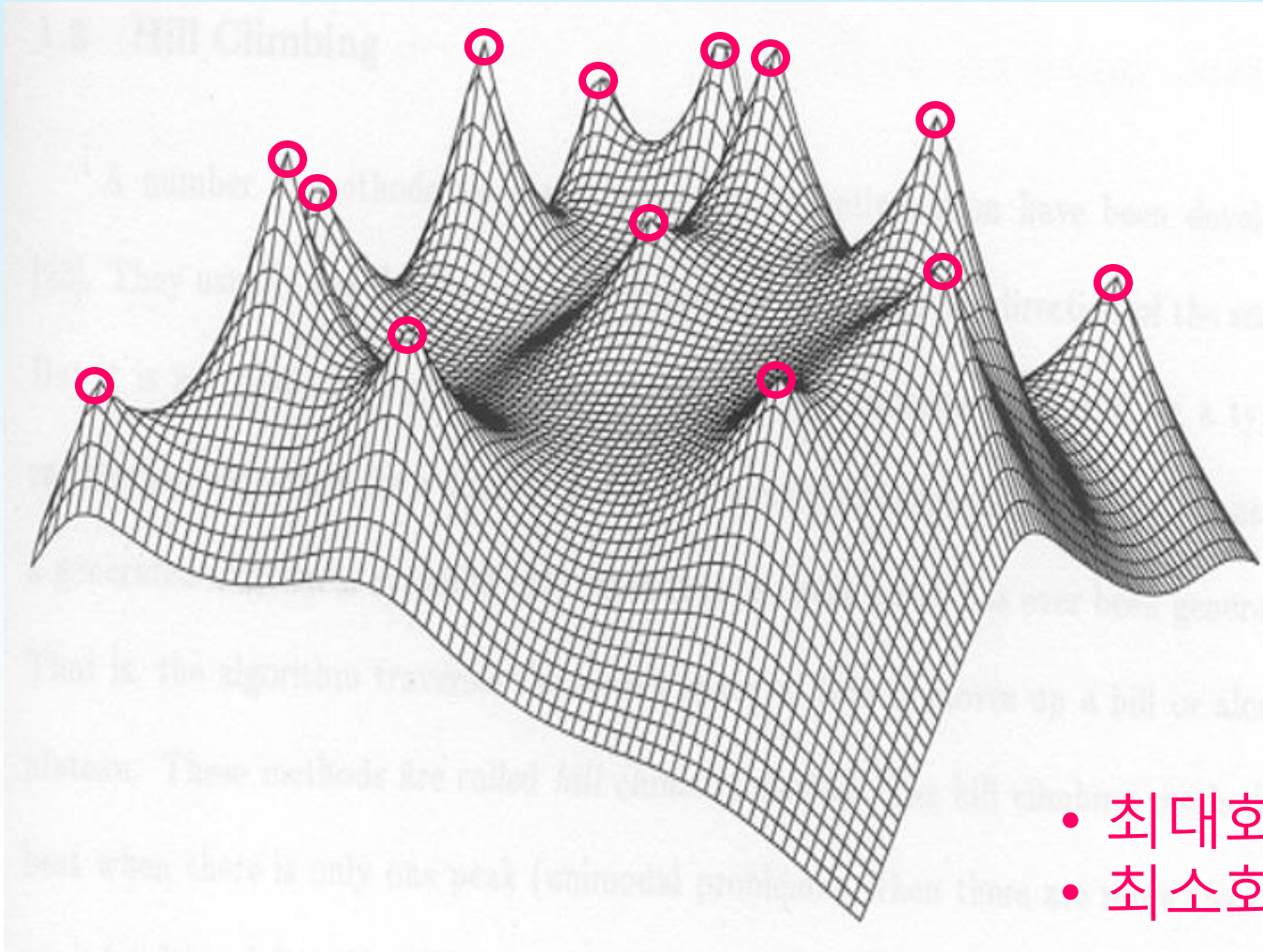
25-city TSP (아주 조그만 문제)

- 25-city problem $\sim 6.2 * 10^{23}$ solutions
- 1만 clocks/solution \rightarrow

모든 경우를 다 보려면 700억년 소요

✓ 8000-city TSP 문제의 공간은 어떨까?

끌개 - Local Optimum



○: 지역최적점
(local optimum)
=
끌개
(attractor)

- 최대화 문제 : 봉우리
- 최소화 문제 : 계곡

- TSP의 끝개 수
 - 10-City TSP: 끝개 평균 4 개
 - 20-City TSP: 끝개 평균 170 개
 - 100-City TSP: 끝개 평균 3.4×10^{16} (3경 4천조)
- TSP의 솔루션의 수
 - 100-City TSP의 모든 솔루션의 수: 9.3×10^{157}
- 100-City TSP에서의 끝개의 비율
 - 3×10^{141} 솔루션당 하나씩의 끝개

✓ 8000-city TSP 문제의 끝개 수는?

✓ 최적화 알고리즘은 이런 공간을 돌아다니는 교통수단이다.

끌개 (Attractor)

- Attractor = 끌개
 - 문제공간 상에서의 지역최적점
 - 공간탐색의 목표이자 장애물



- 끌개의 예
 - 생태계의 종: 개나리, 질경이, 치타, 가젤, ...
 - 인류사의 조직, 제도: 가족, 부족, 국가, 학교, 대통령제, ...
 - 인간의 고정 관념, 사고 체계: 시각, 습관, 편집증, ...
 - 시장에서 정착되는 제품들
 - 타게팅 엔진이 생산하는 솔루션 집합
 - 테니스의 스윙폼



- Optimization은 가장 수준 높은 끌개를 찾는 것
 - 저수준 끌개(low-quality local optimum)에 고착되어 버리지 않도록
 - 다양한 끌개에 접할 수 있도록 넓은 탐색 기능 필요

✓ 우리에게 가장 매력적인 투자전략은 아주 높은 하나의 끌개다!

2. 최적화

문제를 푸는 것

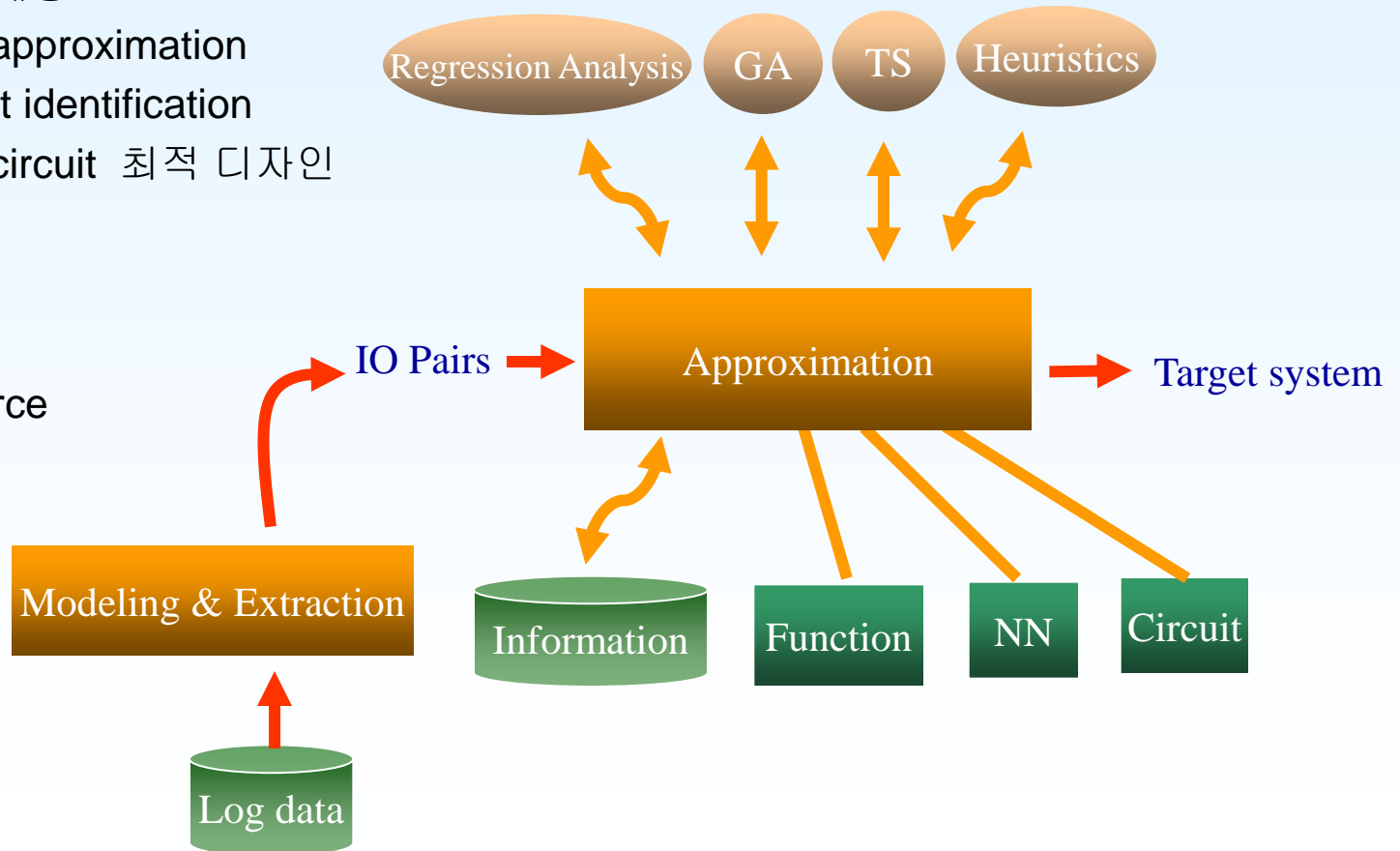
- Decision problems
 - 질문에 대해 Yes 또는 No라고 대답하면 되는 문제
- Equations
 - 조건을 만족하는 해를 찾는 것
 - 유효한 해의 수는 많지 않다 (대부분)
- Optimization
 - 유효한 해는 무수히 많다
 - 이들 중 가장 매력적인 해를 찾는 것

Types of Optimization Problems

- Function optimization
 - 주어진 함수값을 최대화/최소화하는 변수값 찾기
- System identification
 - 주어진 입력-출력 집합을 가장 잘 설명하는 시스템 찾기
 - Function approximation, Neural-network optimization, 투자전략, ...
- Combinatorial optimization
 - 이산적 해들의 집합에서 가장 매력적인 해를 찾는 것
 - 그래프 분할, TSP, 차량 라우팅, VLSI 회로 배치, 벡터 양자화, 전략 최적화, ...

System Identification

- Determines a mathematical model for an unknown system by observing input-output pairs
- Applications
 - 원자력발전소
 - 1-To-1 타게팅
 - Function approximation
 - Neural-net identification
 - Decision circuit 최적 디자인
 - ...
- Areas
 - 금융
 - e-commerce
 - Search
 - ...



- Function optimization
 - 함수의 최적해 탐색, 방정식 근사, ...
- System identification
 - Mostly black-box models
 - » 신경망 디자인, 함수 근사, ...
 - » 문자 인식, 질병 예측, 사기 진단, **주가 예측**, **투자전략**, ...
 - » Decision logic, 다중 엘리베이터 최적제어, 핵연료 상관식 디자인, ...
- Combinatorial optimization
 - TSP, 차량 라우팅, 스케줄링
 - 시재관리 최적화, 공정-장비 할당 최적화
 - 그래프 분할, VLSI 회로 배치
 - 벡터 양자화
 - **투자 전략**, 게임 전략, ...

Optimization Methods

- Deterministic algorithms
- Local search algorithms
 - Greedy algorithms
 - Heuristics
- Linear programming
- Dynamic programming
- Neural net
- Stochastic approaches
 - Genetic algorithm, LSMC, tabu search, ...

- 시간의 제한
 - Hierarchical model 필요
 - 경험적 직관 필요
- Abstraction, articulation, event화
 - Dream: Raw data로부터 바로 시스템 도출
 - » 시스템의 복잡성으로 인해 대부분 가능하지 않다
 - Articulation 필요

3. 주식투자자와 최적화

종합 NEWS 증시 이기는 `마법의 공식`을 찾다
컴퓨터·수학 대가 문병로 서울대 교수
기사입력 2012.11.02 09:11:56 | 최종수정 2012.11.02 09:41:48

기사 나도 한마디 ORACLE 오라클 백서 다운로드 EVENT



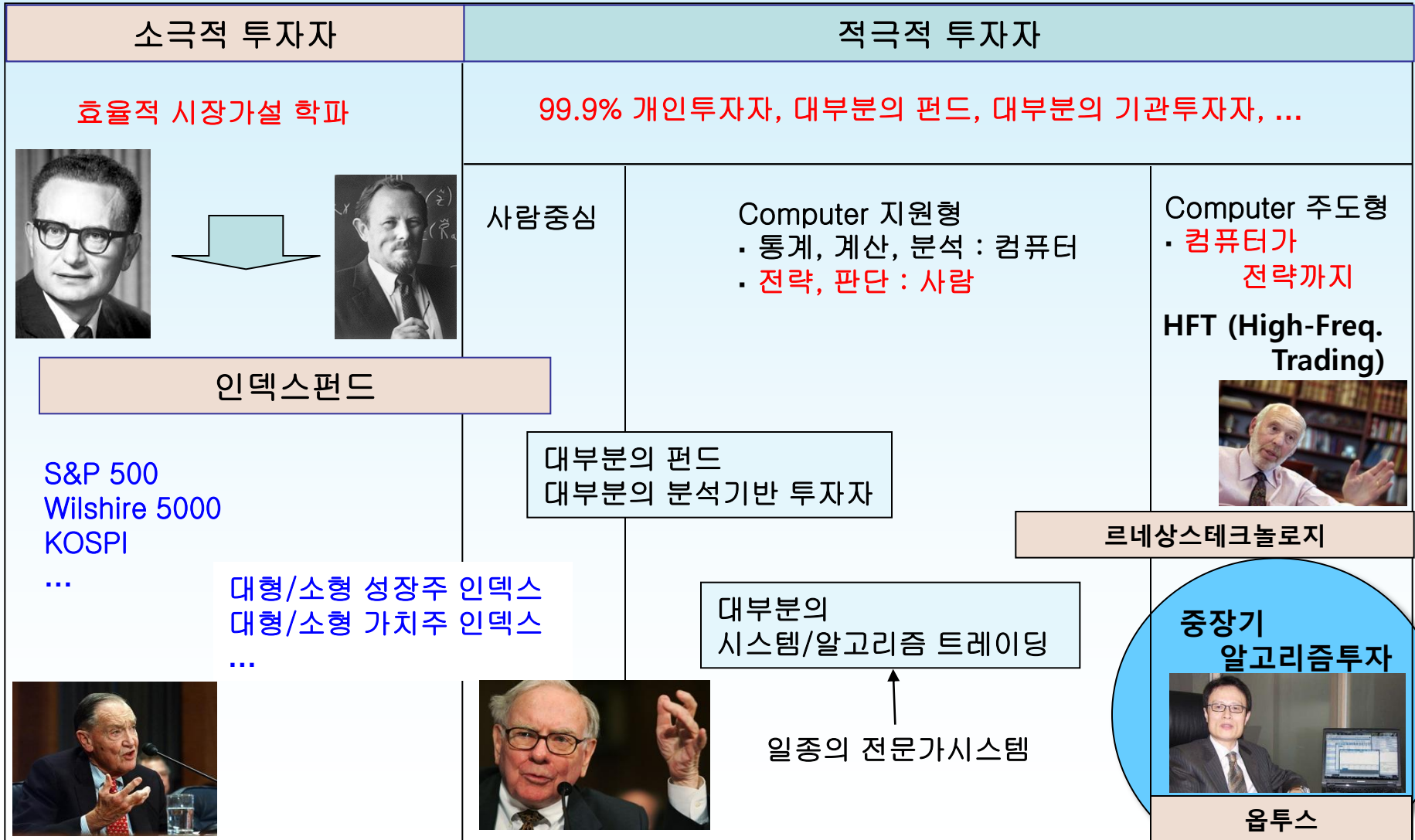
문병로 교수가 알고리즘 이론을 적용한 투자시스템에 대해 도표를 들어 보이며 설명하고 있다. 문 교수는 "주식투자는 본질적으로 리스크를 거머하는 것"이라며 컴퓨터에 의한 주식거래가 시장을 지배하게 될 것이라고 전망했다. <김호영 기자>

‘주식으로 신(神)의 영역에 도전하다.’

주식시장 원리를 한 줄로 설명하는 마법의 공식이 존재한다면 몇 차원의 함수로 풀어낼 수 있을까. 주가에 영향을 미치는 변수는 매우 많다. 최근 주가 움직임을 비롯해 거래량, 대차대조표, 손익계산서, 현금흐름은 물론 환율, 수출입 통계, 세계 증시, 다른 개별종목들의 주가까지. 이들 모든 변수를 고려한 마법의 공식을 만든다는 건 인간의 힘으론 아마 불가능할 것이다. 여기에 천재 수학자이자 컴퓨터 박사가 도전장을 내밀었다. 문병로 서울대 컴퓨터공학부 교수가 그 주인공이다. 문 교수는 컴퓨터 알고리즘 최적화 분야의 세계적인 대가다.

알고리즘이란 고등학교 때 배운 다이어그램이나 순서도를 떠올리면 이해가 비교적 쉽다. 알고리즘을 최적화한다는 것은 컴퓨터를 이용해 어떤 일을 수행하기 위한 가장 효율적인 방법을 찾는 것이다.

투자철학에 따른 분류



알고리즘 트레이딩 현황

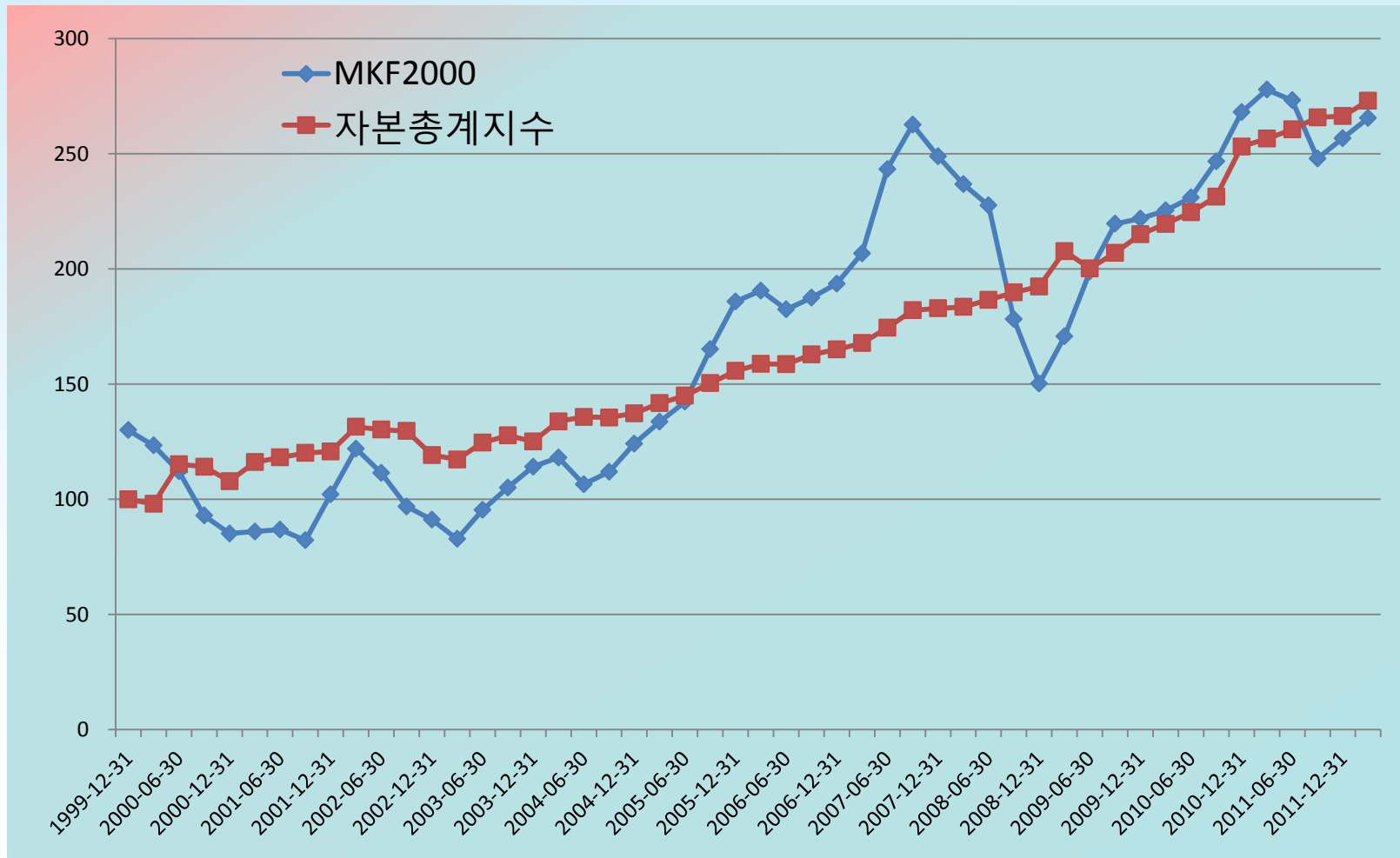
- 2006년의 algorithm trading 비중
 - EU와 미국 주식 거래의 1/3
 - 외환거래 주문의 1/4
 - 런던주식거래소 주문의 40%
- 2010년의 algorithm trading 비중 (예상, Aite Group LLC)
 - EU와 미국 주식 거래의 1/2 (이미 2009년 미국 주식 거래의 73%가 high-frequency trading 회사들의 주문)
 - 옵션 거래의 1/5
- 급격한 변화
 - Avg **“... more people working in their technological area than ... on the trading desk. ... The nature of the market has changed dramatically.”** – 7/2007 AP
 - Trad **... multiplied, it became arms race.** – 11/2006 Herald Tribune
 - Algorithms and quant **“Goldman ... has more people working in their technological area than people on the trading desk. ... The nature of the market has changed dramatically.”** – 7/2007 AP
- ATD
 - 자동 거래 전문 회사
 - 미국 거래의 6% 점유
 - City Group이 6.8억불에 인수
- Nova Fund (Renaissance Technologies)
 - Algorithm trading fund
 - 97년 어느날 NASDAQ 전체 거래의 14% 차지
- 2014년에는
 - 90%↑ ?
- Any sort of pattern recognition or predictive model can be used to initiate algorithm trading. **Neural net** and **genetic programming** have been used to create these models. “Now it’s an **arms race**,” said Andrew Lo, director of the MIT’s Laboratory for Financial Engineering. -Wikipedia



문지기의 꿈 (Janitor's Dream)

- 인지물리학자들
 - 사람의 의식을 표현할 수 있는 수식이 있을 것이다
 - 의식의 추상화 레벨은 10층, 논의는 지하1층
- 일반투자자들, 수학-물리학자들
 - 시장의 원리를 설명하는 마법의 수식이 있을 것이다
 - 시장의 추상화 레벨은 10층, 논의는 1층

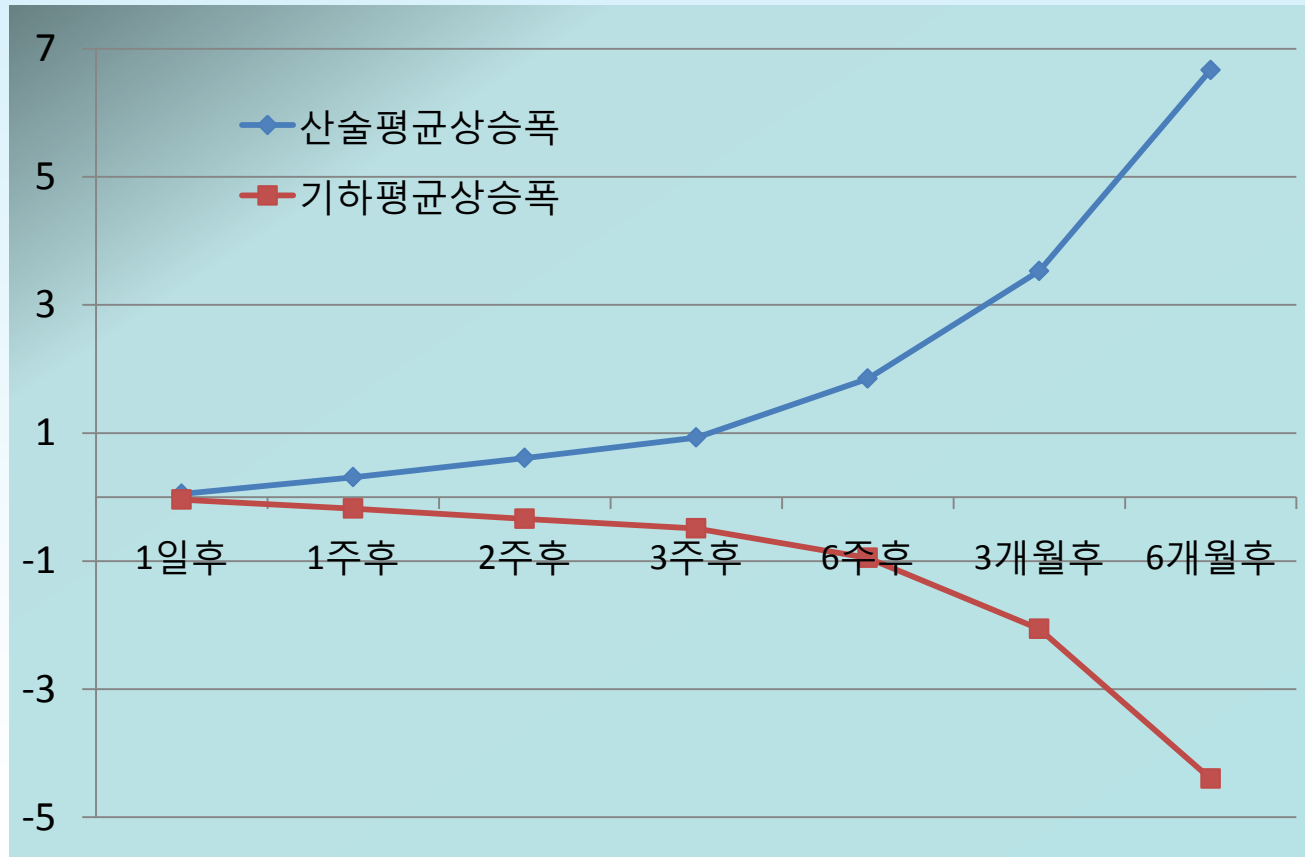
주가는 장기적으로 장부가치를 반영한다



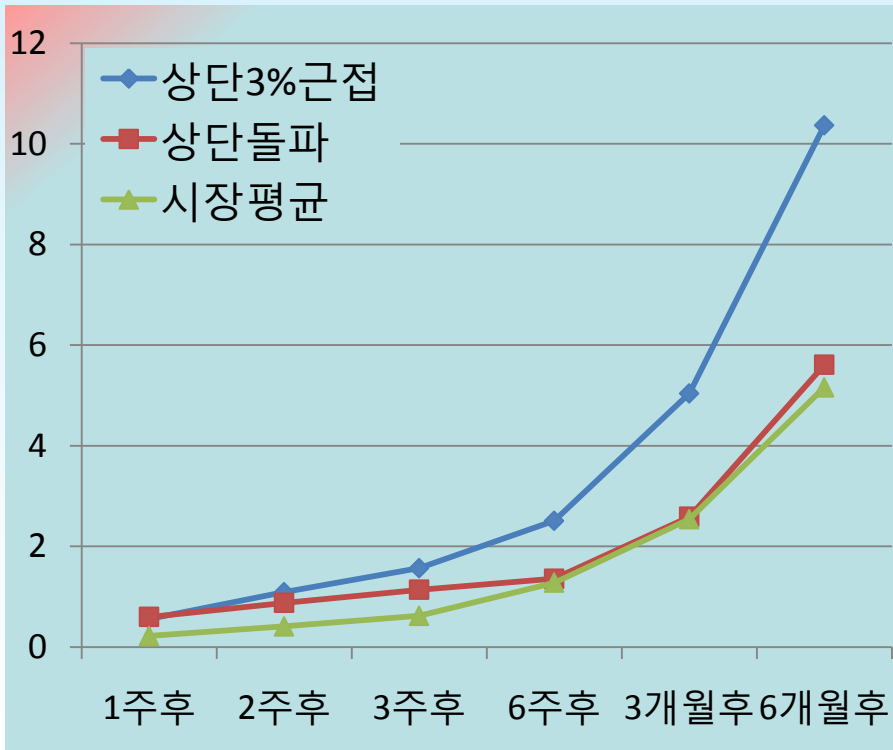
2000년 1월 ~ 2012년 6월

한국증시의 평균상승폭, 2000~2011

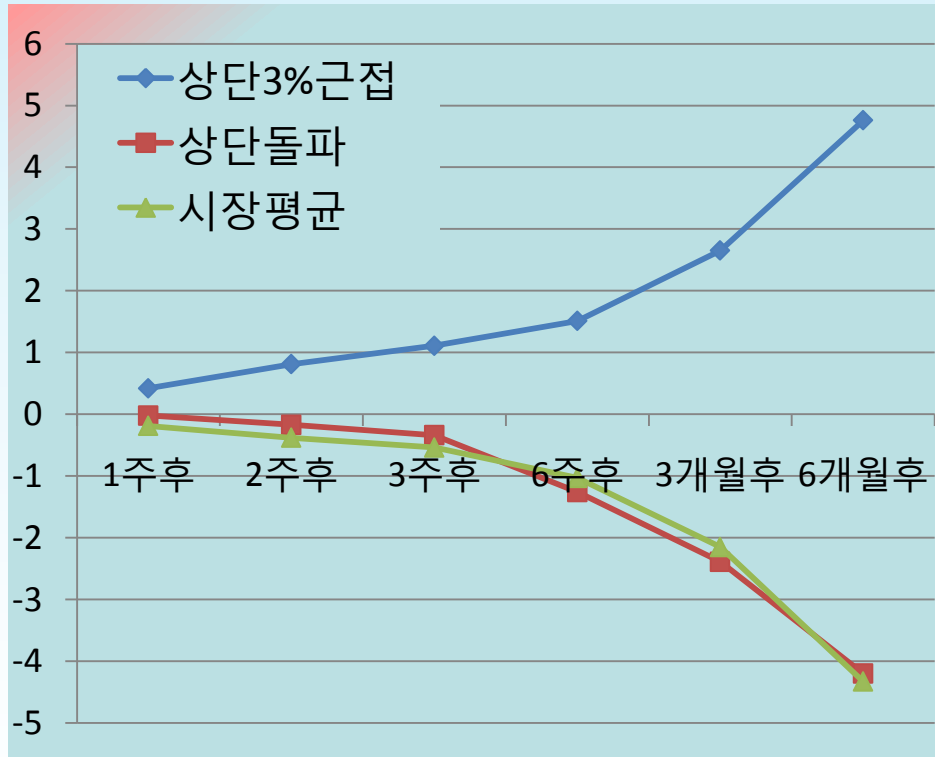
평균 상승폭 (%)



상식과 다른 예: 볼린저밴드



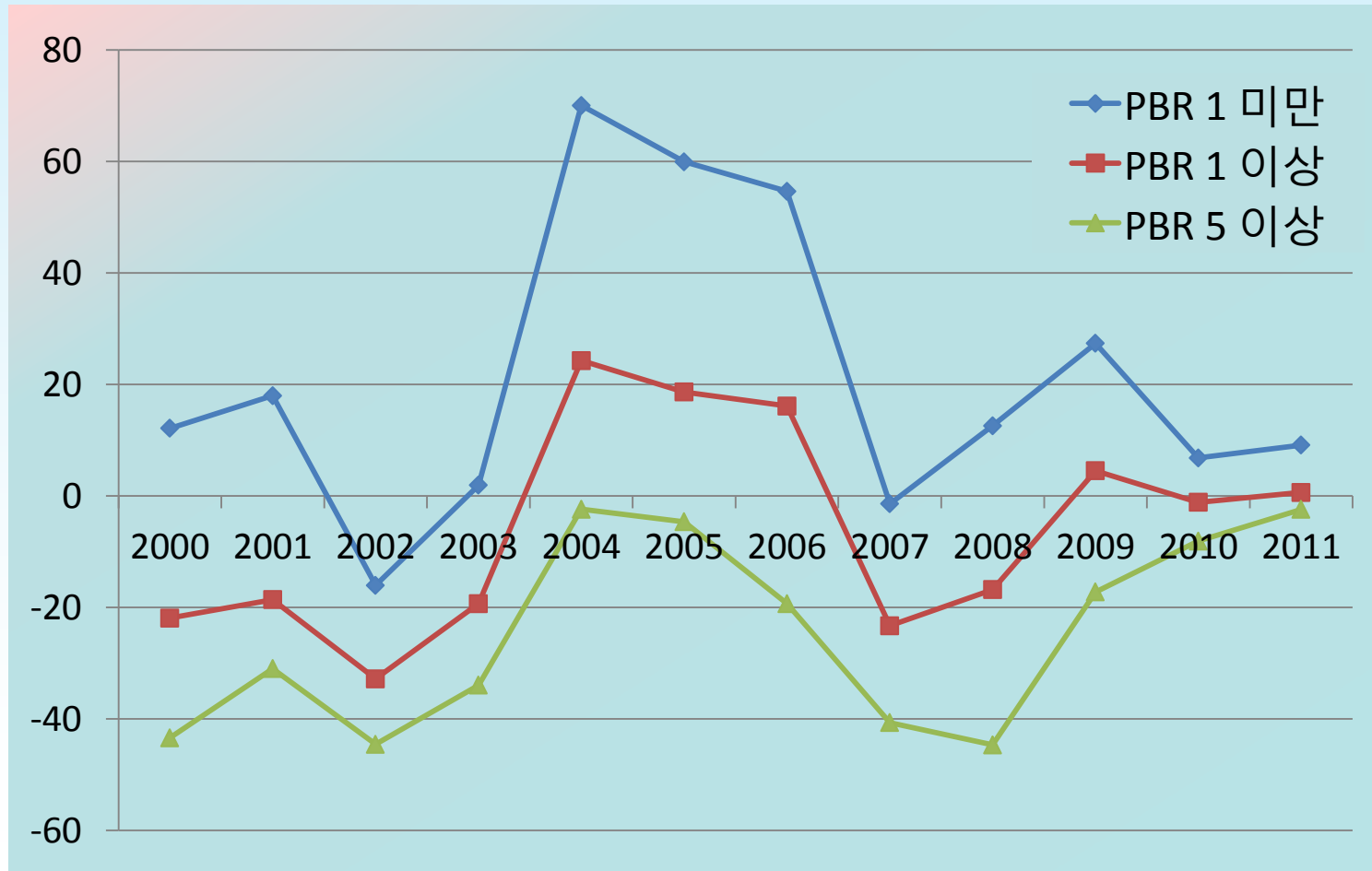
산술평균



기하평균

인기의 실체, w/ PBR

1년 평균 상승폭 (%)



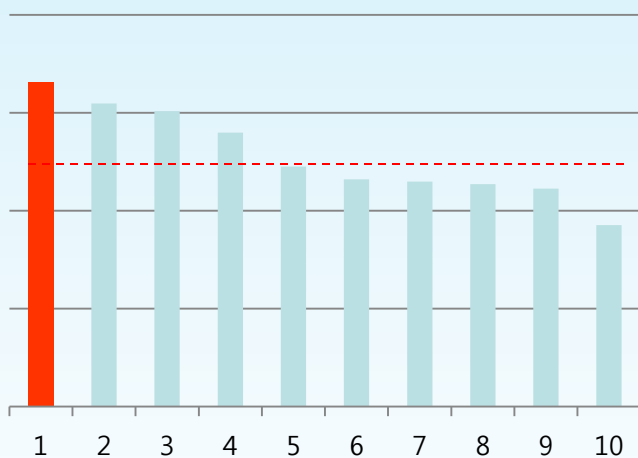
미국 주식의 PBR 십분위에 따른 수익률 1951-2003

비교, 한국 (2000~2012)

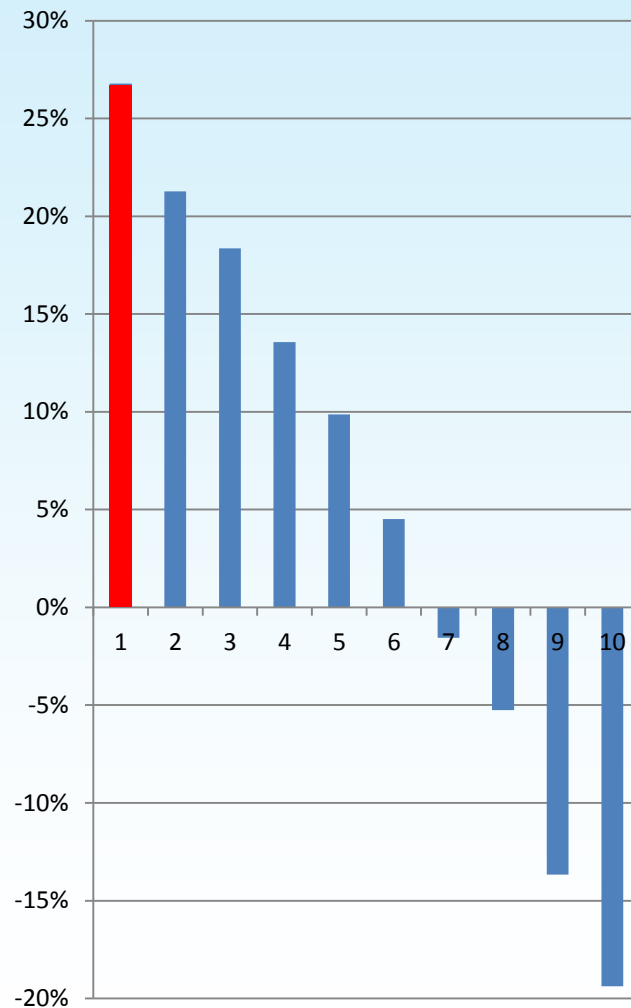
십분위	연평균 기하수익
1 (Lowest)	16.59%
2	15.48%
3	15.08%
4	13.98%
5	12.25%
6	11.60%
7	11.49%
8	11.35%
9	11.12%
10 (Highest)	9.26%

• 동 기간 미국주식
연평균 기하수익률: **11.7%**

기하수익률

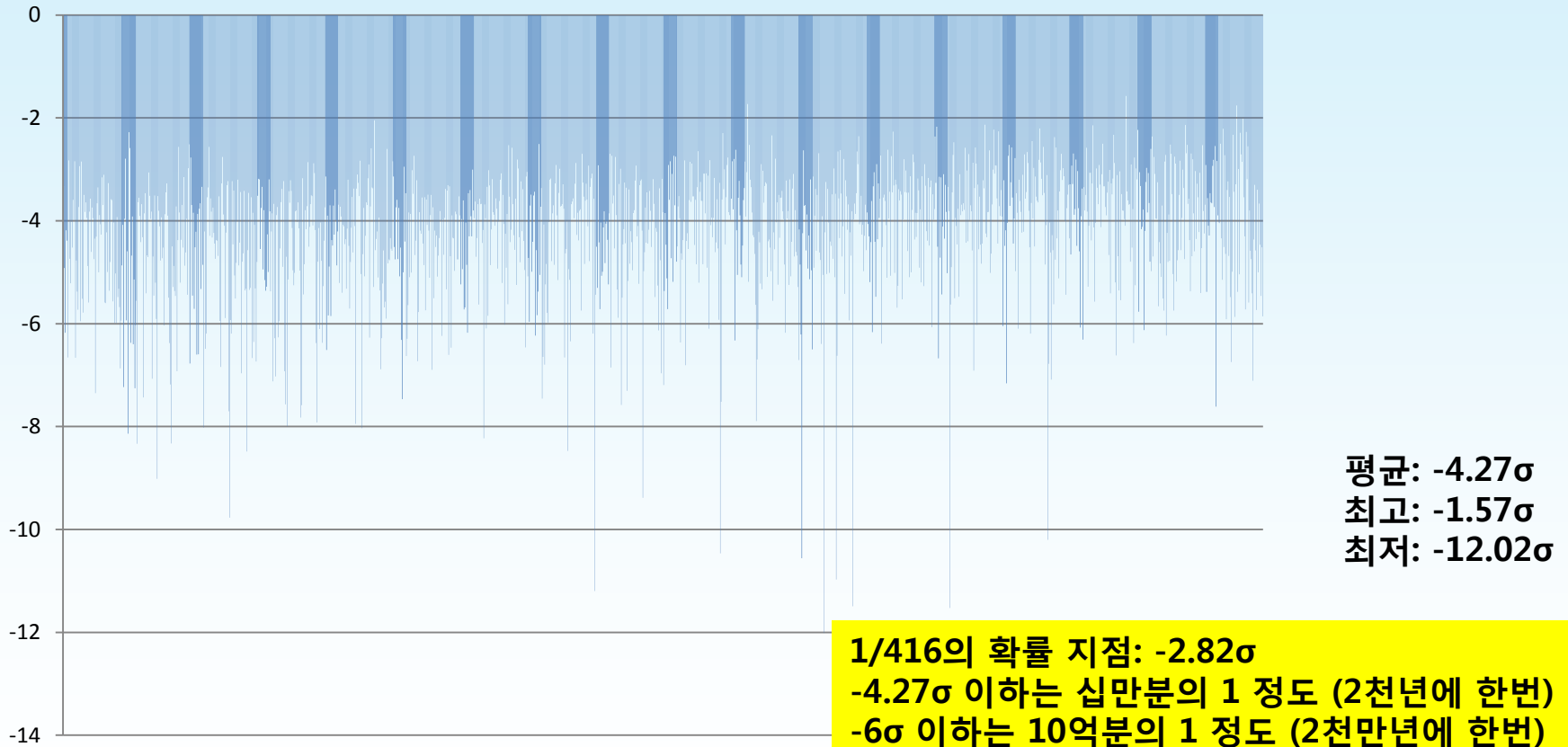


실제 투자수익은 여러 요인으로
이보다 다소 낮게 나온다



극단치의 함정

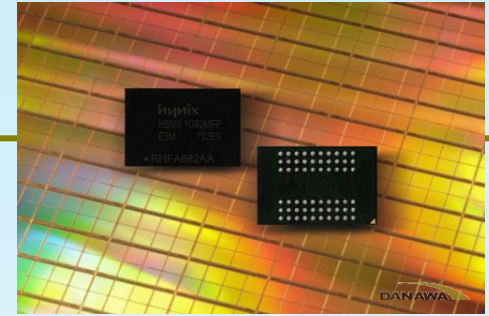
주수익률(Log) 하단 극단치, 8년(전후 4년)



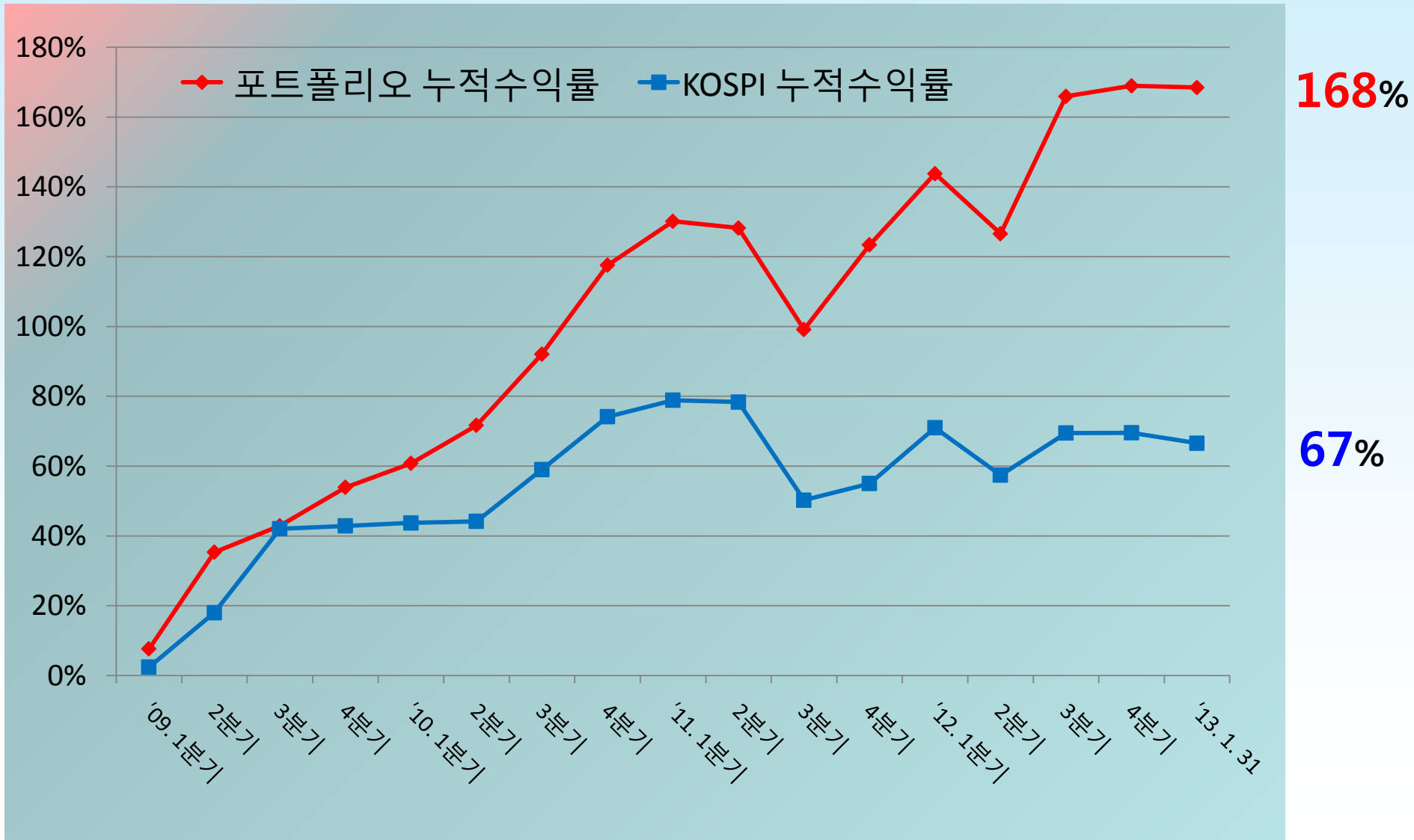
- ✓ 이런 일이 8년 동안 무수히 일어난다
- ✓ 정규분포 턱없이 맞지 않고, log 처리해도 여전히 턱없다

기록들

- Graph partitioning
 - Johnson Benchmark (1994, 2004)
- Circuit partitioning
 - MCNC Benchmark (1998)
- TSP
 - TSP LIB (1994, 2002)
- 198-꼭지 지수귀문도 발견 (2003)
 - 유전 알고리즘 사용
- 행렬곱셈의 복잡도 $O(n^{2.81})$ 알고리즘
 - 유전 알고리즘에 의한 자동발견 (세계 최초)
 - 6700만년 → 10분
- GECCO 2D Assignment Contest 우승 (2008)
- 우리의 관점에서 이런 문제들이나 주식투자 문제 모두 최적화 문제다



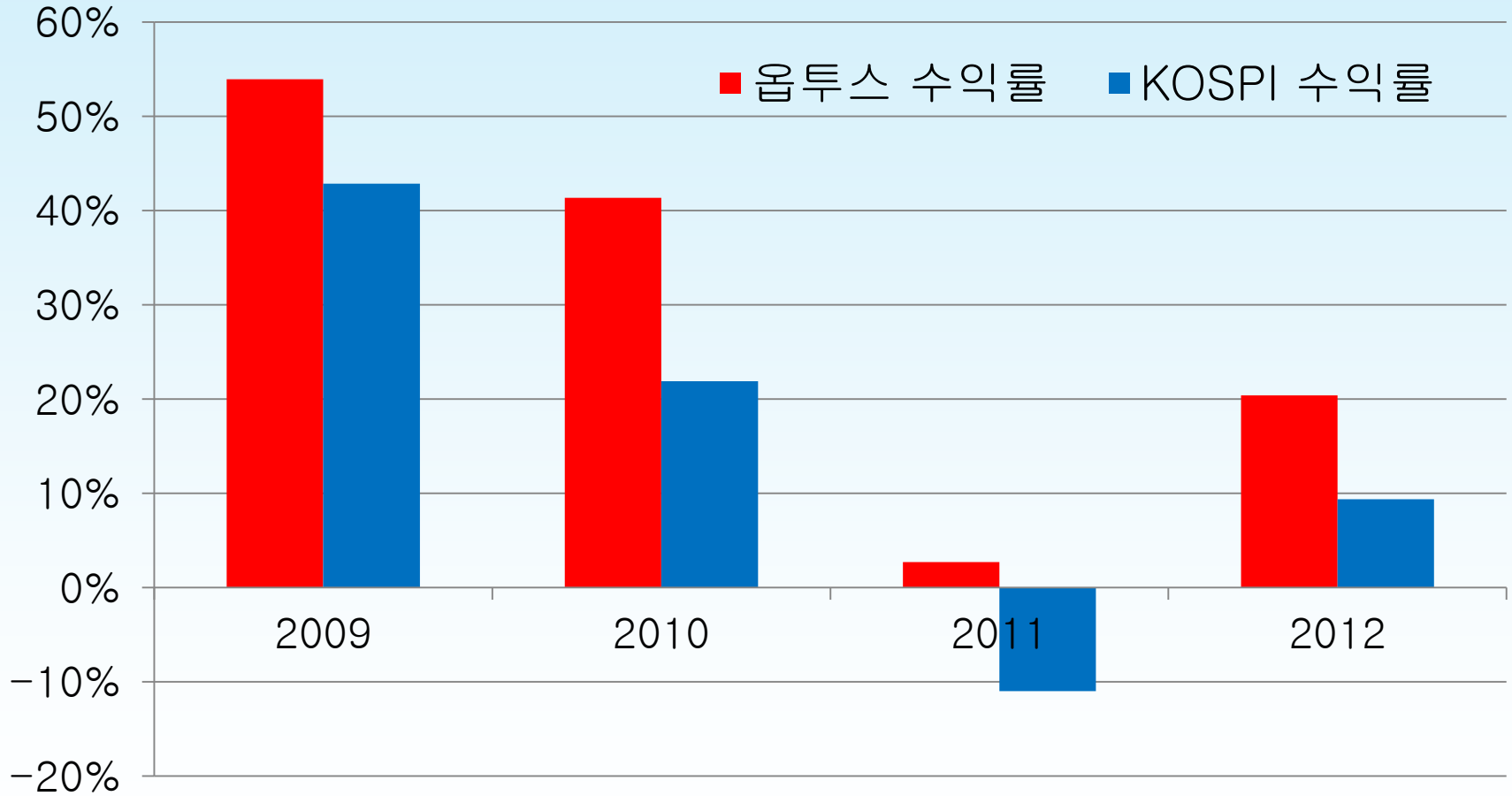
누적 수익률



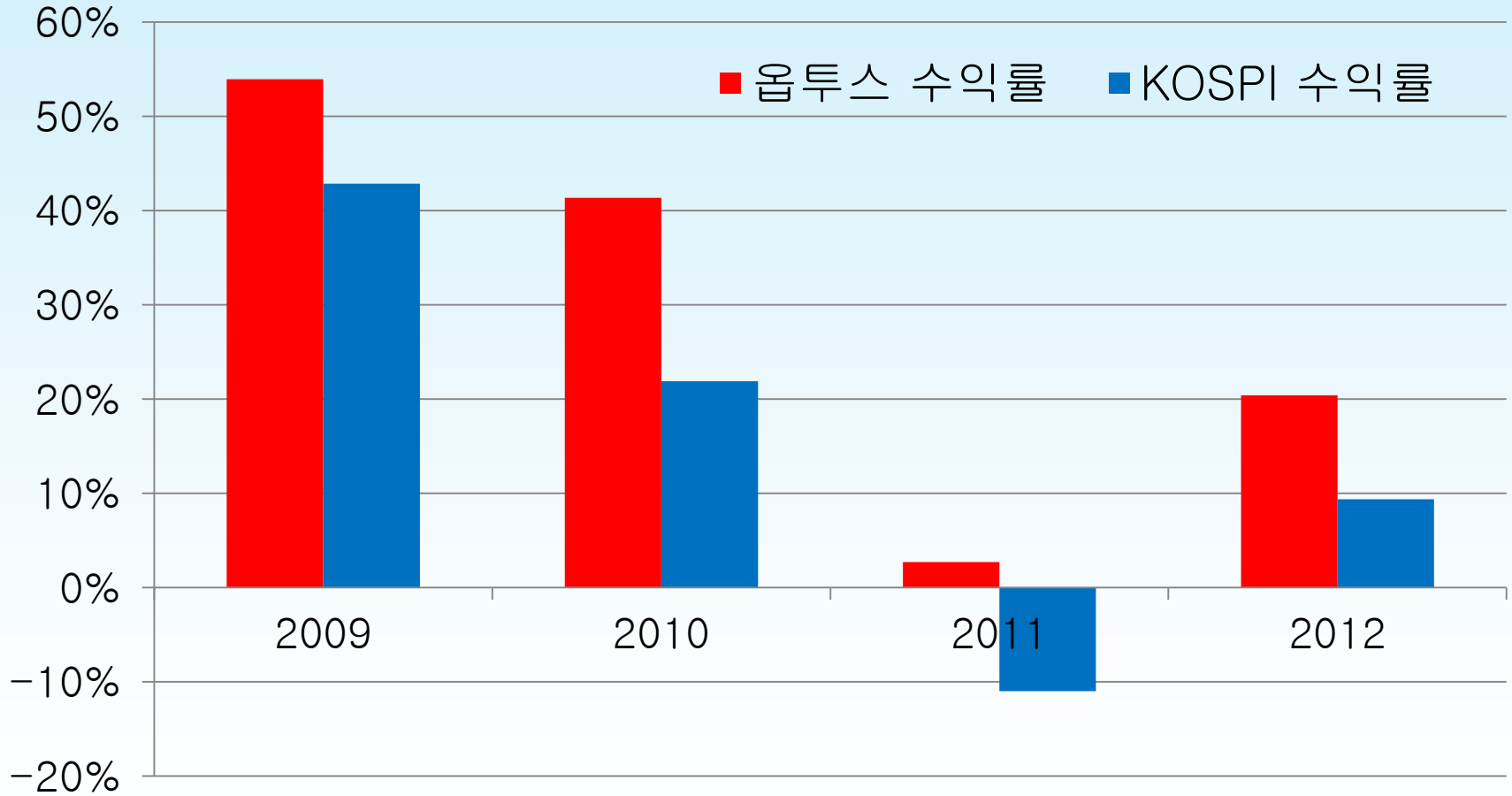
분기	포트폴리오 분기별 수익률	포트폴리오 누적 수익률	KOSPI 분기별 수익률	KOSPI 누적 수익률
'09. 1분기	<u>7.60</u>	7.60%	2.41	2.41%
2분기	<u>25.75</u>	35.31%	15.24	18.01%
3분기	<u>5.60</u>	42.90%	<u>20.36</u>	42.05%
4분기	<u>7.71</u>	53.91%	0.58	42.86%
'10. 1분기	4.45	60.76%	0.60	43.72%
2분기	<u>6.79</u>	71.68%	0.32	44.18%
3분기	11.88	92.06%	10.28	59.00%
4분기	13.26	117.53%	9.51	74.13%
'11. 1분기	5.80	130.14%	2.72	78.86%
2분기	-0.84	128.21%	-0.29	78.34%
3분기	-12.75	99.12%	-15.76	50.24%
4분기	<u>12.19</u>	123.39%	3.17	55.00%
'12. 1분기	9.12	143.77%	10.31	70.99%
2분기	-7.05	126.58%	-7.95	57.40%
3분기	<u>17.37</u>	165.93%	7.67	69.47%
4분기	1.16*	168.96%	0.04	69.55%
'13. 1. 25	-0.19	168.45%	-1.76	66.57%

* 추정 배당 2% 반영

연도별 수익률 비교 with KOSPI



연도별 수익률 비교 with KOSPI



최적화 문제 투성이

- 대차대조표
- 손익계산서
- 현금흐름표
- 주가
- 거래량
- 경제지표

문제 모델링과 최적화 알고리즘

- Portfolio
- 투자전략
- Pattern
- ...

2020년 어느 칼럼...

“10년 전만 해도 이렇게 변수가 많고 복잡한 투자를 기계보다 사람이 더 잘할 수 있다고 대부분 믿었다. 그런 시절도 있었다.”

사람은 아니다

- 사람의 인지적 오류가 발생하는 곳에 수익의 기회가 있다
 - 컴퓨터 지원형까지 인지적 오류가 산재한다
 - 컴퓨터주도형에도 모델링에 인지적 오류가 존재한다
- 주식투자는 한단계 들어가보면 리스크를 거래하는 것이다
- 리스크를 현명하게 견딜 수 있는 사람은 극소수
- 기계적인 매매가 답이다



✓ 당분간은 혼재하면서 점점 컴퓨터 주도형의 비중이 커질 것

결론

- 기능의 시대에서 효율성의 시대로!
- 최적화와 증권시장: 패러다임의 전환
 - 알고리즘들의 전쟁터로 변할 것
 - 확률과 정확도의 게임
- 널린 최적화 문제들
 - 포트폴리오 구성
 - 포트폴리오 리밸런싱
 - 패턴 탐색
 - 투자 전략 최적화
 - 차익거래
 - ...
- 교과서적 이론과 현장의 괴리
 - Customization
 - Variation
 - Abstraction
- Blue ocean!
 - 아는 만큼 보이며, 보는 만큼 이룬다!