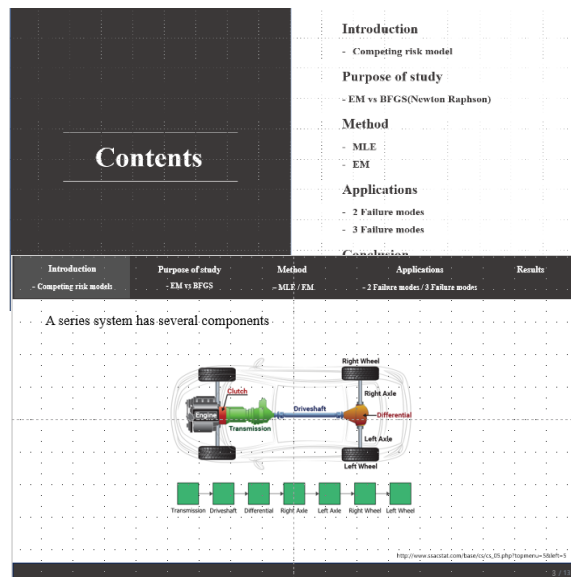


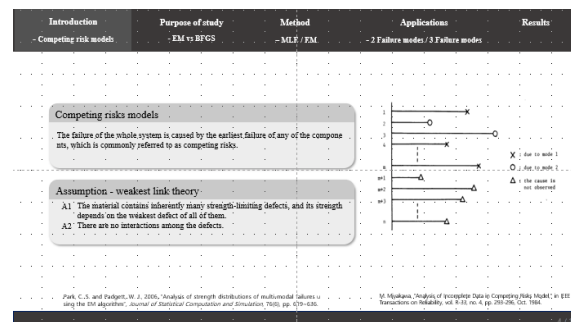
안녕하십니까, 저는 산업공학 응용통계 전공 정나미라고 합니다. 제가 발표할 내용은 ‘EM 알고리즘을 이용한 기계부품의 고장 원인 및 신뢰성 분석’입니다. 발표 시작하겠습니다.



목차는 다음과 같습니다. 경쟁위험모델에 대한 소개 연구목적 방법 소개 어플리케이션 과 결론 순으로 진행하겠습니다. 시간적 여유가 된다면 추후 연구에 대해서도 설명을 드리겠습니다.

다음 그림은 직렬 시스템을 나타낸 것입니다. 자동차를 하나의 시스템으로 볼 때, 이 중 부품 하나라도 고장이 날 경우 자동차는 고장 나게 됩니다. 경쟁위험모델은 이와 같은 직렬 시스템에 사용됩니다.

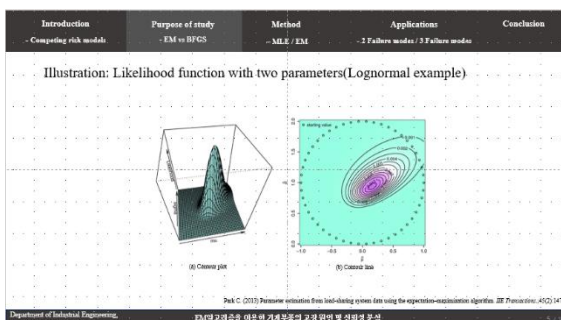
직렬 시스템은 단 하나의 부품이라도 고장이 나면 전체 시스템을 쓸 수 없게 되는 것이 특징입니다.



경쟁위험모델이란 말 그대로 위험을 경쟁한다는 뜻입니다. 시스템의 고장 내에서 부품의 고장끼리 경쟁하여 가장 빠른 고장이 시스템 고장을 불러일으킨다는 것입니다.

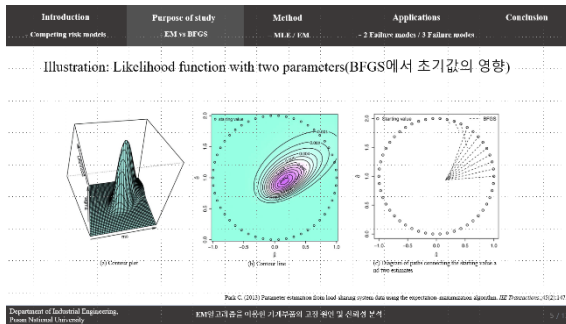
이 모델을 적용하기 위해서는 2가지의 약한 연결 고리 이론 가정이 필요합니다.

고 있으며 시스템의 강도는 그 중 가장 약한 부품의 강도에 달려있다는 것입니다. 다른 하나는 부품에 대한 고장원인이 서로 독립적이라는 것입니다. 다음으로는 경쟁위험모델에 왜 EM



시스템은 다양한 강도를 지닌 부품들을 지니고 있으며 시스템의 강도는 그 중 가장 약한 부품의 강도에 달려있다는 것입니다. 다른 하나는 부품에 대한 고장원인이 서로 독립적이라는 것입니다. 다음으로는 경쟁위험모델에 왜 EM 알고리즘을 적용하게 되었는지 알아보겠습니다. 그림은 우도함수를 최대를 하는/ 로그노말의 최적의 파라미터들을 구하는 시뮬레이션 입니다. 그림(b)를 보시면 파라미터의 초기해인 스타팅 밸류를 50개를 발생시켜 최적의 해로 수렴하는지 알아보는 것입니다. 자세히 보면 보통 closed form의 형태가 아닌, 즉 numerical한 방법으로

MLE를 구할 때 뉴턴-랩슨 계열의 알고리즘을 많이 이용합니다. 그러나 뉴턴랩슨계열의 알고리즘은 최적의 해를 구하는데 초기값의 영향이라는 한계가 있습니다.

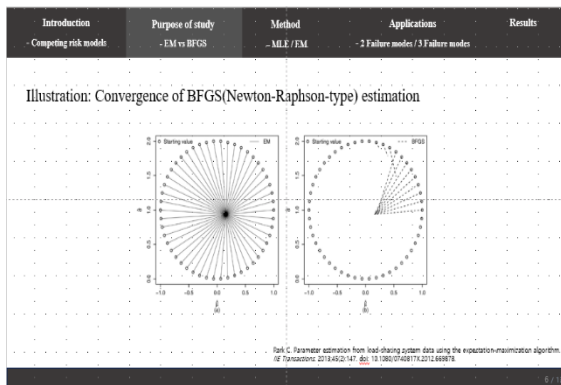


왼쪽은 EM 알고리즘을 쓴 경우이고 오른쪽은 뉴턴 랩슨 계열 중 하나인 BFGS 알고리즘의 결과입니다.

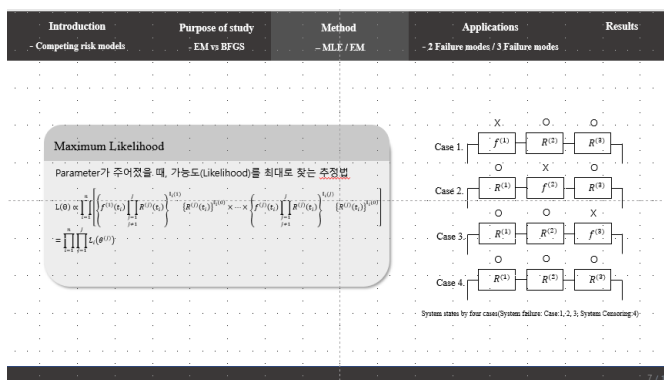
(차이를 간단히 말씀드리면) EM 알고리즘은 수열 방식을 반복 수행하여 최적의 해를 찾는 방식이고 뉴턴 랩슨 계열은 기울기를 이용하여 최적의 해를 갖습니다.

그림 (a)와 (b)는 파라미터별 우도함수의 최대값을 가르키는 등치선인데 이 그림을 반대로

돌려 해를 찾는 과정을 구슬을 떨어뜨려 홀에 들어가는 것이라고 한다면, 기울기에 의해 구멍으로 들어갈 것입니다. 뉴턴 랩슨 알고리즘은 기울기를 이용하여 해를 구하므로 그림 (c)에서 보시다시피 등고선이있는 기울기가 있는 스타팅 밸류에서는 최적의 해를 찾아가지만 약간의 경사에서는 해를 찾아가지 못하는 것을 보실 수 있습니다. 그래서 댄스퍼는 1977년 수열을 반복하는 EM알고리즘을 소개하였습니다. 그리고 그

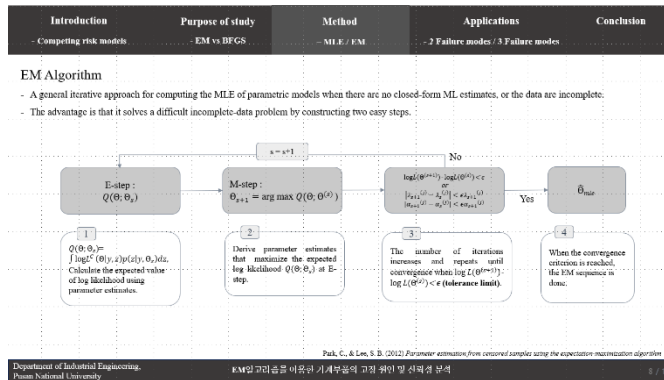


림 (a)처럼 단봉인 유니모달에 대해서는 항상 최대값을 찾는다라는 증명을 하였습니다. 그러나 신뢰성 분석 도구에서는 여전히 EM알고리즘을 쓰지 않고 있었습니다. 따라서 본 연구의 목적은 EM 알고리즘을 이용한 경쟁위험모델을 소개하고 기존 분석 도구와 비교 분석을 하는 것입니다.



EM알고리즘을 설명하기 앞서 EM 알고리즘을 사용하기 위한 우도 함수를 구하는 식입니다. 우측에 그림으로 설명하자면, 4가지의 직렬시스템 내 부품 고장을 나타낸 것입니다. 첫번째 상자를 고장모드1, 두번째 상자를 고장모드2, 세번째 상자를 고장모드3이라고 가정합니다. 첫번째의 경우가 고장 났다면 고장함수의 확률밀도함수를 곱하고 나머지

는 신뢰도 함수를 구하여 곱해줍니다. 두번째의 경우는 고장모드2가 고장난 경우, 세번째는 고장모드 3이 고장난 경우입니다. 네번째의 경우는 센서링 된 경우입니다. 관측 중단이 되었으므로 아무 고장도 나지 않아 모든 부품에 신뢰도 함수를 구하여 곱해줍니다. 이 모든 경우를 곱해주면 우도 함수를 구할 수 있습니다.



이 이전과의 차이가 허용한계치에 수렴하지 않는다면 수렴할 때까지 반복횟수를 증가시킵니다. 수렴 기준에 도달하게 되면 알고리즘은 종료합니다.

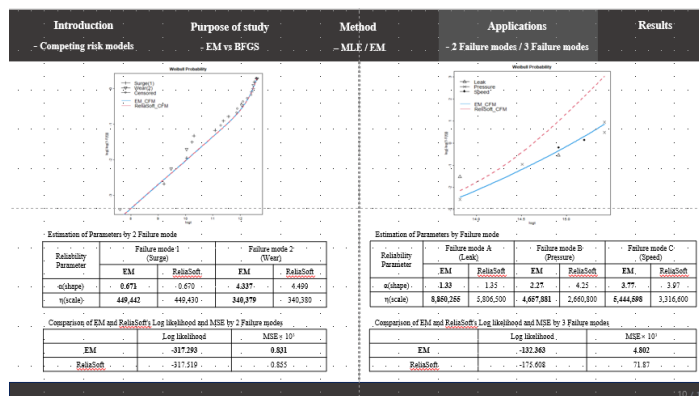
Introduction	Purpose of study	Method	Applications	Conclusion																																																	
- Competing risk models	- EM vs BFGS	- MLE / EM	2 Failure modes / 3 Failure modes																																																		
<div><div>2 Failure modes data 30 observations of devices from a field-tracking study of a large system. - Failure modes : Censoring(0), Surge(1), Wear(2)</div><div>3 Failure modes data - 8 observations of complete data from a pneumatic cylinder(system) - Failure modes : Leak(A), Pressure(B), Speed(C)</div></div>																																																					
<div><div>Competing failure mode: with 2 failure mode and censoring data<table><tr><th>Interval Time (Cycle)</th><td>300,000</td><td>300,000</td><td>2,000</td><td>384,000</td><td>293,000</td><td>88,000</td><td>147,000</td><td>28,000</td><td>143,000</td><td>300,000</td></tr><tr><th>Failure mode</th><td>0</td><td>0</td><td>1</td><td>1</td><td>2</td><td>1</td><td>1</td><td>1</td><td>1</td><td>0</td></tr></table></div><div>Competing failure mode: with 1 failure mode<table><tr><th>Sample Number</th><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td></tr><tr><th>Interval Time (Cycle)</th><td>2,000,000</td><td>3,000,000</td><td>5,000,000</td><td>1,000,000</td><td>4,000,000</td><td>1,000,000</td><td>5,000,000</td><td>3,000,000</td></tr><tr><th>Failure mode</th><td>B</td><td>A</td><td>B</td><td>B</td><td>C</td><td>A</td><td>B</td><td>C</td></tr></table></div></div> <div>Chen, X L S., Chen, D. O., Kwan, B. S., Park, J. W. and Lee, C. S., (2013) Reliability Analysis of a Mechanical Component with Multiple Failure Modes, <i>Proc. Korean Soc. Mech. Eng. A</i>, 37(9), pp. 1169-1174.</div>					Interval Time (Cycle)	300,000	300,000	2,000	384,000	293,000	88,000	147,000	28,000	143,000	300,000	Failure mode	0	0	1	1	2	1	1	1	1	0	Sample Number	1	2	3	4	5	6	7	8	Interval Time (Cycle)	2,000,000	3,000,000	5,000,000	1,000,000	4,000,000	1,000,000	5,000,000	3,000,000	Failure mode	B	A	B	B	C	A	B	C
Interval Time (Cycle)	300,000	300,000	2,000	384,000	293,000	88,000	147,000	28,000	143,000	300,000																																											
Failure mode	0	0	1	1	2	1	1	1	1	0																																											
Sample Number	1	2	3	4	5	6	7	8																																													
Interval Time (Cycle)	2,000,000	3,000,000	5,000,000	1,000,000	4,000,000	1,000,000	5,000,000	3,000,000																																													
Failure mode	B	A	B	B	C	A	B	C																																													
McKee W Q and Escobar L A., (1998) Statistical Methods for Reliability Data, John Wiley & Sons, Inc.																																																					
Department of Industrial Engineering, Pusan National University.																																																					
EM알고리즘을 이용한 기계부품의 고장 원인 및 선택성 분석																																																					
9 / 13																																																					

우도 함수를 구하여 EM알고리즘에 적용시킵니다. EM알고리즘은 2단계가 반복되는 것으로 E스텝에서는 기존의 파라미터 추정량으로 로그우도함수의 기댓값을 계산합니다. M스텝에서는 E스텝에서 구한 로그우도함수의 기댓값을 최대화하는 파라미터를 추정합니다. 그리고 만약 각 모수들의 로그우도함수 값

이러한 방법으로 와이블 분포를 대상으로 2가지 고장모드와 3가지 고장모드일 때의 사례를 적용하였습니다. 2가지 고장모드의 예제는 총 30개 관측치이고 고장모드에는 서지와 마모 2가지 고장모드와 센서링 데이터도 포함되어 있습니다.

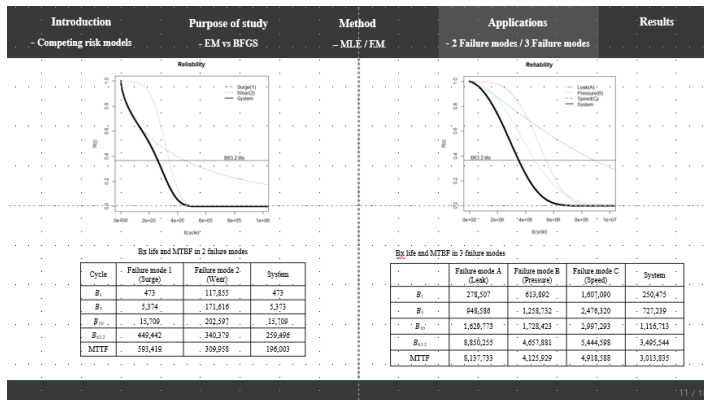
3가지 고장모드 데이터는 아래 논문에서 실험한 공압실린더 데이터입니다

다. 고장모드는 총 3가지로 누수-A, 압력-B, 속도-C로 이루어져 있습니다. 저희는 R을 활용하여 분석하였습니다.

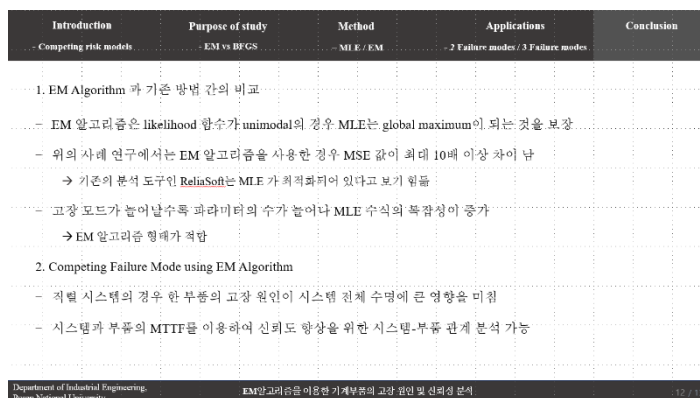


기존 분석 도구인 ReliaSoft의 결과는 아래 논문의 결과 데이터를 가져와 비교하였습니다. CFM은 competing failure mode라는 뜻으로 경쟁위험모델과 같은 뜻이며 릴리아소프트에서 사용되는 약자를 따왔습니다. 좌측은 2개 고장모드 우측은 3개 고장모드의 결과이고, 아래 테이블은 모수 추정 결과 그리고 알고리즘의 정확성

평가를 나타내었습니다. 먼저 2가지 고장모드일 때는 그림상에서는 그래프가 같지만 우도 함수나 MSE값에서는 EM알고리즘이 0.831로 더 나은 결과를 나타내었습니다. 반면에 3가지 고장모드에서는 그래프 상에서도 차이가 보이지만, MSE값은 10배 이상 차이가 나는 것을 알 수 있습니다.



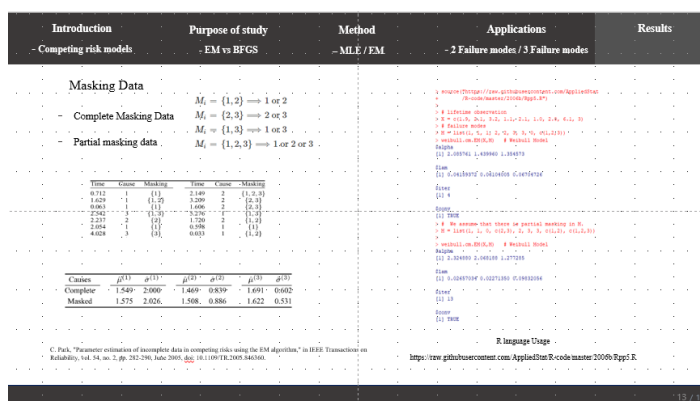
다음은 보수 추정 결과로 신뢰도 그래프를 그려 Bx 수명과 MTTF, 평균고장시간을 구하였습니다. (이 때 시스템의 수명은 신뢰도 역함수를 이용하여 구하였고 MTTF는 평균수명이므로 시스템 그래프의 아래 면적 넓이로 구하였습니다. 직렬 시스템인만큼 처음 고장난 부품 수명이 시스템의 수명과 같은 것을 알 수 있습니다.)



결론입니다. 그래프뉴튼 계열의 방법은 그레디언트 경사하강법에 기초를 하고 있으므로 경사가 거의 없는 곳에서는 해를 찾는 진행이 멈추는 경우가 많다.

(예제 적용을 통해)

(시간 관계 상 추후연구는 생략하고 발표를 마치겠습니다. 감사합니다.)



다음은 주후 연구입니다. 오른쪽 아래에 제가 사용한 EM알고리즘을 이용한 경쟁위험모델의 R 프로그램의 깃허브 주소입니다. 링크에 접속하시면 위 예제를 이용하실 수 있으며, 완전한 데이터, 센서링 데이터 뿐만 아니라 마스킹 데이터도 적용이 가능합니다. 왼쪽 아래의 논문의 예제를 가져왔는데 고장이 났으나 어떤

부품에 의해 고장이 발생했는지 알 수 없을 때를 마스킹 데이터라고 하고 이를 포함한 데이터로 모수를 추정한 결과입니다. 완전한 데이터일 때의 결과 차이가 거의 없다는 것을 알 수 있습니다. 따라서 마스킹 데이터를 포함한 실 데이터를 찾아 연구를 진행해보려고 합니다. 발표 마치겠습니다. 감사합니다.