6주차 읽기자료: Bounds, 완화와 solver 결과

반응기 계산을 검토하는 순서

제품 규격을 만족하는 반응기 조건을 계산했을 때, 최적화가 얼마나 진행됐는지와 공정 예측이 맞는지를 각각 확인합니다.

두 뉴런 예제의 MILP·LP 해를 직접 유도한 뒤 bounds와 scaling을 계산하고, OMLT 연결과 CSTR 결과를 확인합니다.

계산에 필요한 신경망 계수와 실행 결과는 실습 ZIP에 있습니다. LP와 MILP는 모두 오픈소스 HiGHS로 풉니다.

Big-M에서 삼각형 완화 유도

$L < 0 < U$이고 $z \in [0, 1]$이라고 놓습니다. $h \le Uz$와 $h \le a - L(1 - z)$를 z에 대해 정리합니다.

$$\frac{h}{U} \le z \le \frac{h - a + L}{L}$$

$$h\,(U - L) \le U\,(a - L)$$

L이 음수이므로 나눌 때 부등호가 뒤집힙니다. 여기에 $h \ge 0$과 $h \ge a$를 유지하면 강의의 삼각형 완화가 됩니다. 유효한 a의 구간도 유지합니다.

두 뉴런 예제의 정수해 유도

$$0 \le x_1, x_2 \le 1, x_1 + x_2 = 1$$
$$h_1 = \max\left(0, x_1 + x_2 - 1\right), h_2 = \max\left(0, x_1 - x_2\right)$$
$$\min_{x_1, x_2} x_1 \text{subject to} h_1 + h_2 \ge .4$$

입력 등식에서 $x_2 = 1 - x_1$이므로 $h_1 = 0$이고 $h_2 = \max(0, 2x_1 - 1)$입니다. $h_2 \ge 0.4$를 만족하려면 활성 분기에서 $2x_1 - 1 \ge 0.4$여야 합니다.

최적점은 $(x_1, x_2) = (0.7, 0.3)$, 출력은 $(h_1, h_2) = (0, 0.4)$입니다. 목적값 0.7을 두 MILP 구현과 대조합니다.

두 LP 완화의 목적값 계산

구간 bounds [−1, 1]을 사용한 첫 뉴런의 삼각형은 $a_1 = 0$에서도 $h_1 \le 0.5$를 허용합니다. 따라서 $x_1 = 0$, $x_2 = 1$, $h_1 = 0.4$, $h_2 = 0$은 원래 LP의 가능점입니다. 비음수 x_1을 최소화하므로 최적값은 0입니다.

등식에서 얻은 $a_1 = 0$을 bounds에 반영하면 $h_1 = 0$입니다. 두 번째 뉴런의 삼각형 상한은 $h_2 \le x_1$이므로 $x_1 \ge 0.4$가 필요합니다.

이때 $x_1 = 0.4$, $x_2 = 0.6$, $h_2 = 0.4$인 점은 강화한 LP의 가능점이므로 최적값은 0.4입니다. 이 점의 실제 ReLU 출력 h_2는 0이라 정수해가 될 수 없습니다.

완화 해와 실제 신경망의 대조

각 방식에서 목적값 x_1을 최소화합니다. $x_2 = 1 - x_1$이고 필요한 출력 합은 0.4입니다.

방식	x_1, x_2	허용된 h_1, h_2	실제 ReLU 출력	해석
상자 bounds LP	0,1	0.4,0	0,0	가짜 h_1로 규격 충족
등식 반영 LP	0.4,0.6	0,0.4	0,0	가짜 h_2로 규격 충족
MILP	0.7,0.3	0,0.4	0,0.4	원래 출력과 일치

등식 반영 뒤 두 번째 삼각형의 상한은 $h_2 \le x_1$입니다. 따라서 $h_2 \ge 0.4$를 만족하려면 $x_1 \ge 0.4$입니다. 그러나 $x_1 = 0.4$에서 실제 $a_2 = -0.2$이므로 ReLU 출력은 0입니다. 하한이 좋아져도 완화가 원래 그래프와 같아진 것은 아닙니다.

두 뉴런 예제의 분기와 탐색 종료

$h_1=0$을 반영한 문제에서 h_2의 이진변수 z_2만 남깁니다. $L=-1, U=1$이고 $a_2=2x_1-1$입니다. 루트 LP의 $h_2\le z_2$, $h_2\le 2x_1-z_2$를 더하면 $h_2\le x_1$입니다. $h_2\ge 0.4$이므로 $x_1\ge 0.4$가 하한입니다.

문제	계산	결론
루트 LP, $0\le z_2\le 1$	$x_1=0.4, h_2=0.4, z_2=0.4$ 가능	하한 0.4, 정수해 아님
$z_2=0$인 자식	$h_2=0$과 $h_2\ge 0.4$ 충돌	불가능, 종료
$z_2=1$인 자식	$h_2=2x_1-1\ge 0.4$	$x_1=0.7$에서 최소 달성

두 자식은 가능한 이진값을 모두 덮습니다. 한쪽은 불가능하고 다른 쪽은 최소 0.7이므로 전역 하한도 0.7입니다. 가능해 비용 0.7과 같아져 gap=0으로 종료합니다.

여러 열린 노드가 남으면 그중 가장 작은 하한이 아직 더 좋은 해를 허용하는 가지를 나타냅니다. 하한이 incumbent 이상인 가지는 개선할 수 없어 제외합니다. 실제 solver의 분기 순서와 cut은 다를 수 있습니다.

구간 전파 손계산

$a = 2x_1 - 3x_2 + 0.5$, $x_1 \in [0, 1]$, $x_2 \in [-1, 2]$를 계산합니다. 양의 계수는 같은 방향의 극값을, 음의 계수는 반대쪽 극값을 선택합니다.

$$L = 2 \cdot 0 - 3 \cdot 2 + .5 = -5.5$$
$$U = 2 \cdot 1 - 3 \cdot (-1) + .5 = 5.5$$

ReLU 출력 범위는 [0, 5.5]입니다. 이 입력 상자에서는 양쪽 끝점이 실제로 가능하지만 입력 간 등식이 추가되면 같은 계산이 느슨해질 수 있습니다.

LP-OBBT bounds의 유효성

S를 정확한 신경망의 가능집합, P를 $S \subseteq P$인 LP 완화라고 놓습니다. 더 큰 집합에서 최소화하면 값이 작아지거나 같고, 최대화하면 커지거나 같습니다.

$$\min_P a_j \le \min_S a_j$$
$$\max_S a_j \le \max_P a_j$$

따라서 P에서 얻은 두 값은 S의 유효한 바깥 bounds입니다. 기존 구간 bounds와 교집합을 취해 적용합니다. 실습의 P에는 제품 규격이 들어 있지 않아 규격 변경에도 같은 입력 상자에서 사용할 수 있습니다.

미완료 bound 계산의 방향

a를 최소화하는 계산에서 찾은 가능점 값은 실제 최소값의 상한입니다. 이를 a의 하한으로 사용하면 유효한 신경망 점을 잘라낼 수 있으므로 최소값에 대한 유효한 하한을 사용해야 합니다.

a를 최대화할 때는 반대로 최대값의 유효한 상한이 필요합니다. 최대화 가능점 값만으로 상한을 만들면 안 됩니다.

제공 구현은 각 LP의 optimal 종료를 요구하고 `best_objective_bound`를 사용합니다. 바깥쪽으로 1e-7 × (1 + abs(bound))만큼 여유를 두지만, 이 수치 여유가 정확 산술의 증명을 대신하지는 않습니다.

예를 들어 a의 최소값을 구하다가 가능점 a=−0.2와 유효 하한 −0.6에서 멈췄다면 실제 최소는 [−0.6,−0.2]에 있습니다. 뉴런의 하한에는 −0.6을 사용합니다. −0.2를 사용하면 실제 a=−0.4인 점을 잘라낼 수 있습니다. 최대값 계산에서는 반대로 유효 상한을 사용합니다.

뉴런 고정과 영역 변경

a = $x_1 + x_2 - 1$은 입력 상자 [0, 1] × [0, 1]에서 부호가 바뀝니다. 여기에 $x_1 + x_2 = 1$이 있으면 항상 0이므로 h = 0으로 고정합니다.

이 등식을 제거하면서 고정된 h = 0을 유지하면 정식화가 틀립니다. 예를 들어 (1, 1)에서 실제 h는 1입니다.

Bounds 파일에는 입력 범위, 적용한 추가 제약, 고정 가중치의 버전을 함께 보관합니다. 새 모델을 학습하거나 영역을 늘리면 bounds를 다시 계산합니다.

Scaling된 입력과 제품 규격

$$x_1 = \frac{T - 310}{40}, x_2 = \frac{\tau - 2}{8}$$
$$T = 310 + 40x_1, \tau = 2 + 8x_2$$

T = 330 K, τ = 6 min이면 (x_1, x_2) = (0.5, 0.5)입니다. OMLT의 `OffsetScaling`에 입력 offset [310, 2], factor [40, 8]을 전달합니다.

출력이 s = (B − 0.3)/0.1로 표준화되어 있다면 B ≥ 0.38은 s ≥ 0.8입니다. 실습은 출력 factor 1과 offset 0을 사용하므로 물리 농도와 OMLT 출력이 같습니다.

목적함수 변환과 gap 계산

$$J = \tau + \alpha\left(T - 310\right), \alpha = .02\text{min/K}$$
$$J = 2 + 8x_2 + .8x_1$$

J의 incumbent가 3.2, 하한이 2.8이면 절대 gap은 0.4이고 상대 gap은 12.5%입니다. 상수 2를 뺀 식에서는 같은 두 값이 1.2와 0.8이 되어 상대 gap은 약 33.3%입니다.

최적 입력과 절대 gap은 같지만 상대 gap은 달라집니다. 보고하는 목적함수, 상수항, solver의 gap 정의를 같이 적습니다.

목적함수 상수와 gap의 계산

동일한 가능해와 하한에서 목적함수의 상수 2만 제거한다고 둡니다.

계산	상수 포함 J	상수 제거 J−2
가능해 목적값	3.2	1.2
유효 하한	2.8	0.8
절대 gap	3.2−2.8=0.4	1.2−0.8=0.4
상대 gap	0.4/3.2=12.5%	0.4/1.2≈33.3%

모든 후보의 값에서 같은 수를 빼므로 후보 순위와 최적 입력은 유지됩니다. 하지만 상대 gap의 분모는 달라집니다. 두 보고서를 비교할 때 목적함수 정의가 다르면 %만으로 풀이 품질을 비교할 수 없습니다.

가능해가 없으면 분모에 사용할 값도 없습니다. 그때는 유효 하한과 가능해 부재를 보고하고, 변수의 초기값을 해로 해석하지 않습니다.

OMLT의 입력 정보 구성

2입력 모델의 scaling과 입력 bounds를 설정합니다. 양의 factor로 변환한 입력 구간은 둘 다 [0, 1]입니다.

```
scaling = OffsetScaling(
    offset_inputs=[310., 2.], factor_inputs=[40., 8.],
    offset_outputs=[0., 0., 0.],
    factor_outputs=[1., 1., 1.])
definition = NetworkDefinition(
    scaling_object=scaling,
    scaled_input_bounds={0: (0., 1.), 1: (0., 1.)})
```

각 `DenseLayer`에 가중치와 bias를 넣고 `add_edge`로 연결합니다. 실습 배열의 가중치 축은 [입력, 출력]이며 입력 순서는 [온도, 체류시간]입니다.

고정 입력 검사 절차

1. 입력 영역의 모서리와 내부점을 선택합니다.
2. 원래 신경망의 forward 출력을 계산합니다.
3. 같은 입력을 MILP에 고정하고 제품 규격을 제외합니다.
4. 가능해 존재와 종료 상태를 확인한 뒤 출력을 읽습니다.
5. 출력 차이, 등식·부등식 잔차와 정수성 잔차를 비교합니다.

정식화 검사는 제품 규격을 제외한 신경망 제약만으로 수행합니다. 제품 규격을 남기면 규격을 위반하는 입력에서 검사 문제가 infeasible이 됩니다. 실습은 최대 출력 차이 1e-5 mol/L 미만을 요구합니다.

부등식 잔차는 만족할 때 0이 되도록 계산합니다. 예를 들어 $B \geq 0.38$의 위반량은 $\max(0,0.38-B)$, $C \leq 0.25$는 $\max(0,C-0.25)$입니다. 이진값 $z=0.999999$의 정수성 잔차는 가까운 정수 1과의 거리 0.000001이며, solver 허용오차와 비교합니다.

기준 CSTR의 수지와 속도식

정상상태·등온·일정 밀도에서 A → B → C의 연속 1차 반응을 가정하고 유입 A 농도를 1.2 mol/L로 둡니다.

$$C_A = \frac{1.2}{1 + k_1\tau}$$

$$C_B = \frac{k_1\tau C_A}{1 + k_2\tau}, C_C = k_2\tau C_B$$

330 K에서 k_1 = 0.2, k_2 = 0.1 /min입니다. 온도 의존성은 다음 페이지의 교육용 Arrhenius 계수로 정합니다. 이 수지는 기준 모델이며 신경망 근사와 별도로 계산합니다.

온도 의존성과 데이터 생성

$$k_i\left(T\right) = k_i\left(330\right)\exp\left(\frac{E_i}{R}\left(\frac{1}{330} - \frac{1}{T}\right)\right)$$

활성화에너지는 E_1 = 40,000, E_2 = 50,000 J/mol, 기체상수 R = 8.314 J/(mol·K)입니다. 이 값은 예제에 지정한 매개변수로, 실제 물질에서 추정하지 않았습니다.

입력 상자에서 균일 난수로 학습점 1,200개를 생성합니다. 학습 seed는 61이고, 독립 검사·test 생성에는 seed 87을 사용합니다. 구현은 test 데이터를 계수 적합에 사용하지 않습니다.

학습 설정과 고정 계수 사용

2 → 12 → 12 → 3 구조에서 앞의 두 affine layer 뒤에 ReLU를 적용하고 마지막 층은 선형으로 둡니다. Loss는 물리 단위로 나타낸 세 농도 출력의 평균제곱오차입니다.

Full-batch Adam을 6,000회 적용하고 학습률은 처음 3,999회 0.006, 나머지는 0.002입니다. $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\varepsilon = 1e\text{-}8$을 사용합니다.

기본 실행은 `frozen_cstr_network.json`의 계수를 사용합니다. `--refit` 결과는 지정한 출력 폴더의 별도 파일로 저장되며 기본 계수를 덮어쓰지 않습니다. 비교 결과는 사용한 계수 파일과 함께 보관합니다.

동일 문제의 계산 결과 읽기

정식화	LP 하한	MILP 최적값	이진변수
loose	2.0000	3.1656	24
interval	2.0000	3.1656	22
LP-OBBT	2.5215	3.1656	19

OMLT도 같은 MILP 최적값을 반환했습니다. 입력·가중치·제품 규격·목적함수는 네 실행에서 같습니다.

표의 LP 하한은 이진변수를 연속으로 바꾼 별도 문제의 최적값입니다. Solver가 presolve와 cut을 적용한 뒤 보고하는 bound와는 다를 수 있습니다.

계산 시간의 측정 범위와 bounds 재사용

정식화	Bounds [s]	구성 [s]	MILP [s]	합계 [s]
loose	0.0001	0.0016	0.3107	0.3125
interval	0.0001	0.0016	0.1539	0.1556
LP-OBBT	0.1132	0.0016	0.0955	0.2104

합계는 bounds, 모델 구성과 MILP 풀이의 합입니다. 별도로 측정한 LP 완화, 예측 검증, 학습 시간은 포함하지 않습니다.

고정 모델과 영역에서 bounds를 재사용하면 여러 풀이에 전처리 비용을 나눌 수 있습니다. 반복 비교에서는 같은 장비·스레드·시간 제한을 사용하고 중앙값과 변동을 기록합니다.

bounds를 재사용할 때의 총시간

기록된 한 번의 실행에서 구간 전파의 MILP 시간은 0.1539 s, LP-OBBT는 0.0955 s입니다. OBBT 전처리는 0.1132 s가 필요했습니다.

풀이 수 m	구간 전파: 0.0001+m×0.1555	OBBT: 0.1132+m×0.0971
1	0.1556 s	0.2103 s
2	0.3111 s	0.3074 s
5	0.7776 s	0.5987 s

열별 1회 비용에는 build 0.0016 s를 포함합니다. 표는 반올림된 기록 시간을 사용하고 각 반복의 풀이 시간이 같다고 가정한 산술 계산입니다. 실제 반복 실험 결과는 아닙니다.

고정된 가중치와 입력 영역에서 bounds를 재사용할 수 있을 때 전처리 비용을 나누어 부담합니다. 새 모델이나 더 넓은 영역이면 bounds 유효성을 다시 확인해야 하며 이 계산을 그대로 적용할 수 없습니다.

연습: scaling과 gap

T = 340 K, τ = 4 min을 scaled 입력으로 바꾸고 목적함수 J를 계산합니다. 이어서 incumbent 5.0, 하한 4.5의 절대·상대 gap을 계산합니다.

확인: (x_1, x_2) = (0.75, 0.25), J = 4 + 0.02 × 30 = 4.6입니다. Gap은 0.5와 10%입니다.

Incumbent 없이 하한 4.5만 있으면 무엇을 보고할까요? 확인: 유효한 하한과 가능해 미확보를 보고하며, 상대 gap과 선택 입력은 비워둡니다.

연습: bounds와 기준 공정 검사

입력 상자에서 계산한 bounds를 사용하다가 온도 상한을 350 K에서 360 K로 늘렸습니다. 기존 bounds를 그대로 적용해도 될까요? 확인: 확장된 영역에서 bounds를 다시 계산하거나 유효성을 증명해야 합니다.

선택 입력의 신경망 B는 0.380000, 기준 CSTR B는 0.371157 mol/L입니다. 규격 여유를 계산합니다. 확인: 기준값에서 하한을 뺀 여유는 -0.008843 mol/L로 음수입니다.

이 경우 solver gap을 줄이는 것만으로 규격 위반을 고칠 수는 없습니다. 예측 모델 또는 선택 조건을 다시 검토해야 합니다.

원문과 실습 파일

OMLT 논문: Ceccon 등, JMLR 23(349), 2022 · 공식 사용 예제

Anderson 등, Strong mixed-integer programming formulations for trained neural networks, Mathematical Programming, 2020

HiGHS 공식 사이트와 MIT 라이선스 안내 · Pyomo의 HiGHS 인터페이스

`results.json`은 종료 상태와 검증값, `comparison.csv`는 비교표, `frozen_network.json`은 해당 실행의 계수를 담습니다. 검증 환경은 Python 3.12.14, NumPy 2.5.4, Pyomo 6.10.1, OMLT 1.2.2, HiGHS 1.15.1입니다.

강의노트 · 실습 ZIP