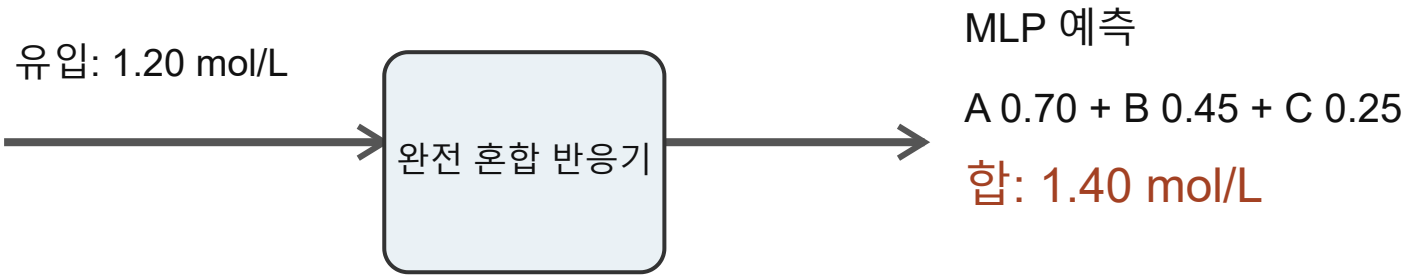
4주차: 물리 제약을 반영한 학습

예측 농도 합과 유입 농도의 불일치

유입 농도가 1.20 mol/L인 반응기에서 MLP가 A·B·C를 [0.70, 0.45, 0.25] mol/L로 예측했습니다. 합은 1.40입니다. 이 출력을 다음 공정의 입력으로 보내도 될까요?



0.20 mol/L는 어디서 생겼을까?

연결 공정에서의 물질수지 오차 전파

다음 분리기는 각 성분의 유입량으로 부하를 계산합니다. 반응기 surrogate의 합 오차가 그 계산에 그대로 들어가므로, 예측 오차와 물질수지 위반을 따로 확인해야 합니다.

이번 주의 주제: loss의 위반 벌점, 지정한 선형 등식으로의 출력 투영, 보정 전후의 성분 오차

물리 제약 학습의 비교 항목

1. 같은 데이터로 MLP·soft sum penalty·hard sum projection을 비교
2. 최소 거리 투영을 KKT 조건으로 유도하고 rank 가정을 확인합니다.
3. 고정 투영층의 역전파를 손계산과 자동미분으로 대조합니다.
4. 성분별 오차·등식 위반·음수 농도·미반영 수지를 따로 보고

실습 기준 모델: 2주차의 정상상태 CSTR

기준 공정과 단위

완전 혼합·일정 부피·같은 유입 및 유출 유량을 가정합니다. A만 유입되며 1:1 연속반응 A → B → C가 일어납니다. 온도 T는 주어지고 에너지수지는 모델에 포함하지 않습니다.

$$C_A = \frac{C_{Af}}{1 + k_1\tau}$$

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

330 K에서 $k_1 = 0.2\ \text{min}^{-1}$, $k_2 = 0.1\ \text{min}^{-1}$입니다. $\tau = 5$ min, C_Af = 1.2 mol/L이면 [C_A, C_B, C_C] = [0.6, 0.4, 0.2] mol/L입니다.

정상상태 수지와 농도 합 등식

각 정상상태 성분수지의 우변을 r_A, r_B, r_C라고 놓고 더하면 반응항이 상쇄됩니다.

$$r_A + r_B + r_C = \frac{C_{\mathrm{Af}} - \left(C_A + C_B + C_C\right)}{\tau}$$

정상상태에서는 각 residual이 0이므로 S = C_A + C_B + C_C = C_Af입니다. 이 농도 합 보존은 위 화학양론과 유량 가정에 의존합니다. 일반 반응에서는 원소수지나 질량수지를 해당 단위로 다시 구성해야 합니다.

A가 B로 반응한 양은 A 수지에서 음수, B 수지에서 양수로 한 번씩 나타납니다. B에서 C로 넘어간 양도 상쇄됩니다. 그래서 합 수지에는 외부 유입과 유출만 남으며, 이 상쇄가 농도 합 제약의 근거입니다.

성분별 예측 오차와 농도 합 오차

예측 [mol/L]	성분 오차 [mol/L]	합 residual [mol/L]
[0.70, 0.45, 0.25]	[+0.10, +0.05, +0.05]	+0.20
[0.70, 0.30, 0.20]	[+0.10, −0.10, 0.00]	0.00

기준값은 [0.60, 0.40, 0.20] mol/L입니다. 두 번째 예측은 합을 맞추지만 A와 B가 틀립니다. 등식 만족만으로 예측 품질을 판단하면 이 오차를 놓칩니다.

물리 residual의 정의

농도 합 residual은 r_sum = C_A + C_B + C_C − C_Af이며 단위는 mol/L입니다. 알려진 속도식을 사용하면 성분별 정상상태 residual도 계산할 수 있습니다.

$$r_A = \frac{C_{Af} - C_A}{\tau} - k_1 C_A$$

$$r_B = k_1 C_A - \frac{C_B}{\tau} - k_2 C_B$$

$$r_C = k_2 C_B - \frac{C_C}{\tau}$$

세 residual의 단위: mol/(L·min). 학습 대상: 합 residual. 별도 평가: 성분수지 residual

총농도가 맞는 [0.7,0.3,0.2]를 330 K, τ=5에 넣으면 종별 잔차는 [−0.04,0.05,−0.01] mol/(L·min)입니다. 합은 0이지만 각 종의 정상상태 수지는 맞지 않습니다. 하나의 합 제약이 세 종별 식을 대체하지 못합니다.

Soft penalty의 데이터 오차항과 제약 위반항

$$L_{\text{soft}} = \frac{1}{3N}\sum_i \left\| \frac{\hat{y}_i - y_i}{c_*} \right\|^2 + \lambda \frac{1}{N}\sum_i \left(\frac{r_{\text{sum},i}}{c_*} \right)^2$$

두 항은 c_* = 1 mol/L로 무차원화합니다. 실습의 λ = 10은 합 residual 한 개와 성분 오차 세 개의 평균 정의까지 포함한 선택

λ를 바꾸면 학습 목표와 gradient 크기가 달라집니다. 유한한 λ와 학습 종료만으로 모든 입력의 residual이 0이라고 보장할 수는 없습니다.

예측을 데이터 쪽으로 움직이는 기울기와 합을 맞추는 기울기가 동시에 작용합니다. 두 요구가 다른 방향을 가리키면 유한한 λ에서는 절충점이 생깁니다. λ의 숫자만 비교하지 말고 각 항을 몇 개의 값으로 평균했는지도 맞춥니다.

유한한 벌점의 해를 직접 구하기

스칼라 예제로 ½(z − a)² + ½λ(z − b)²를 최소화. 데이터는 a를, 등식은 z = b를 요구하며 λ ≥ 0. 미분을 0으로 놓으면:

$$z_* = \frac{a + \lambda b}{1 + \lambda}$$

$$z_* - b = \frac{a - b}{1 + \lambda}$$

a ≠ b일 때 residual은 유한한 λ에서 남습니다. λ를 매우 크게 잡으면 손실항의 크기 차이와 수치 조건도 확인해야 합니다.

미분식을 모으면 (1+λ)z=a+λb입니다. a=1.4,b=1.2,λ=9에서는 10z=12.2로 z=1.22를 얻습니다. 데이터 방향 미분은 −0.18, 벌점 방향 미분은 9×0.02=+0.18로 서로 상쇄되지만 등식 residual 0.02는 남습니다.

자동미분의 계산 대상

현재 실습의 MLP는 [T, τ, C_Af]에서 정상 농도로 가는 mapping입니다. Loss를 파라미터 θ로 미분하며, 합 residual과 정상상태 성분수지에는 시간 미분이 필요하지 않습니다.

연속시간 PINN에서 y_θ(t)를 사용한다면 residual은 $\partial y_\theta/\partial t - F(y_\theta, u)$처럼 좌표 미분도 포함합니다. 이는 θ에 대한 학습 gradient와 구분해야 합니다.

시간을 $\tilde{t} = (t - \mu_t)/s_t$로 표준화했다면 $\partial y/\partial t = (1/s_t)\partial y/\partial \tilde{t}$

Hard 선형 등식과 가능한 출력 집합

$$Ax + By = b,$$
$$By = d(x), d(x) = b - Ax$$

$x \in \mathbb{R}^p$, $y \in \mathbb{R}^n$, $A \in \mathbb{R}^{mxp}$, $B \in \mathbb{R}^{mxn}$, $b \in \mathbb{R}^m$. A·B·b는 고정 계수이고, raw 출력은 $\hat{y} = f_\theta(x)$

CSTR에서는 $x = [T, \tau, C_{Af}]$, $A = [0, 0, -1]$, $B = [1, 1, 1]$, $b = 0$입니다. 따라서 $d(x) = C_{Af}$입니다.

입력 x를 고정하면 d(x)도 숫자로 정해집니다. 합 제약의 경우 출력 세 개를 모두 정하지 않고 합 하나만 정하므로, 합을 유지하면서 A와 B를 서로 바꾸는 자유가 남습니다.

선형 등식 집합으로의 최소 거리 투영

$$y_p = \text{argmin}_z \frac{1}{2} \left\| z - \hat{y} \right\|^2 \text{subject to} Bz = d(x)$$

고정값: 입력 x와 raw 예측 ŷ. 조정 대상: 출력 z. 거리 기준: 물리 농도 좌표의 Euclidean 거리

목적함수는 strictly convex입니다. 가능한 집합이 비어 있지 않으면 투영 출력은 유일합니다. 등식만 있는 이 작은 QP는 KKT 조건으로 닫힌 형태를 구할 수 있습니다.

이미 가까운 예측을 크게 바꾸지 않으면서 합을 맞추려는 선택입니다. 세 농도를 같은 단위와 가중치로 비교하면 수정량 제곱합을 최소화합니다. 오차가 작은 성분을 더 보호하려면 별도의 가중 거리 정의가 필요합니다.

Lagrangian과 KKT 조건

$$\mathcal{L}(z,\lambda) = \frac{1}{2}\left\|z-\hat{y}\right\|^2 + \lambda^T(Bz-d)$$

정지조건과 primal feasibility를 함께 만족시킵니다.

$$z-\hat{y}+B^T\lambda = 0, Bz = d$$

부등식이 없으므로 complementary slackness나 multiplier의 부호 제약은 없습니다. Strictly convex 목적함수와 affine 등식에서는 이 조건을 만족하는 출력이 전역 최소점입니다.

정상성 식은 수정 벡터가 제약면의 법선 B^T 방향이라는 뜻입니다. 제약면을 따라 움직이는 성분은 합을 고치지 못하면서 거리만 늘리므로 최소 거리 해에는 남지 않습니다.

Multiplier를 제거해 투영식 얻기

정지조건에서 $z = \hat{y} - B^T\lambda$를 얻고 $Bz = d$에 대입합니다.

$$G = BB^T, G\lambda = B\hat{y} - d$$
$$y_p = \hat{y} - B^T G^{-1}\left(B\hat{y} - d\right)$$

rank(B) = m이면 $G = BB^T$는 양의 정부호입니다. 식에서는 역행렬로 표현하지만 구현은 $G\lambda = B\hat{y} - d$를 풉니다. 고정 B라면 분해나 고정 투영 계수를 재사용할 수 있습니다.

미지수 λ는 제약 개수 m만큼입니다. 먼저 원래 잔차 $B\hat{y}-d$를 구하고, G에 대한 선형계를 풀어 잔차를 지울 수정량을 계산합니다. B^T를 곱하면 제약 공간의 보정 계수가 출력 n개에 대한 수정 벡터가 됩니다.

KKT-hPINN의 고정 affine 출력층

$$P = I - B^T G^{-1} B$$
$$y_p = P f_\theta(x) + B^T G^{-1}(b - Ax)$$

이 층의 계수는 물리 등식으로 정하고 학습하지 않습니다. MLP 가중치 θ는 투영된 출력의 loss로 학습합니다.

$$By_p = B\hat{y} - GG^{-1}\left(B\hat{y} - d\right) = d$$

지정한 등식은 학습 중과 추론 시 모두 만족합니다. 실제 구현에서는 floating-point 오차와 선형계의 수치 조건을 확인합니다. 출처: Chen et al. (2024).

제약 행렬의 rank와 등식계의 일관성

경우	판단
rank(B) = m ≤ n	모든 d에 출력 해가 존재하며 BB^T가 가역입니다.
중복 행, 일관된 d	가능한 출력은 존재해도 BB^T는 특이합니다. 독립 행이나 SVD를 사용합니다.
중복 행, 모순된 d	가능한 출력이 없습니다. Pseudoinverse로도 모든 등식을 맞출 수 없습니다.

거의 종속인 행은 rank 판정 tolerance와 condition number도 살펴봅니다. 등식의 단위와 크기를 맞추는 row scaling은 가능한 집합을 바꾸지 않습니다.

농도 합 투영을 손으로 계산하기

$$B = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}, d = 1.2$$

$$y_p = \begin{bmatrix} .70 \\ .45 \\ .25 \end{bmatrix} - \frac{.20}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} .633333 \\ .383333 \\ .183333 \end{bmatrix}$$

세 성분에서 같은 0.066667 mol/L를 뺍니다. 합은 1.20이지만 기준 [0.60, 0.40, 0.20]과는 다릅니다.

값은 표시 자릿수로 반올림했습니다. 계산에서는 중간값을 반올림하지 않고, 최종 합 residual을 확인합니다.

합 잔차에서 투영된 출력까지

원시 출력은 [0.70,0.45,0.25], 목표 합은 1.20 mol/L입니다. 세 수정량을 Δ로 같게 두면 1.40−3Δ=1.20이므로 Δ=0.20/3입니다.

단계	값
Bŷ−d	1.40−1.20=0.20
G=BB^{T}	$1^2+1^2+1^2=3$
Gλ=잔차	λ=0.0666667
출력별 수정	[−0.0666667,−0.0666667,−0.0666667]
보정 출력	[0.6333333,0.3833333,0.1833333]

균등 보정의 거리 제곱은 $3\times(0.20/3)^2=0.013333$입니다. A에서만 0.20을 빼도 합은 맞지만 거리 제곱은 0.04로 더 큽니다. 같은 가중치의 Euclidean 거리에서 균등 보정이 선택되는 이유입니다.

독립 등식 두 개의 투영 계산

출력 합 1.2에 더해 A−B=0.2를 요구하는 선형 예제입니다. B의 행을 [1,1,1]과 [1,−1,0], 우변을 [1.2,0.2]로 두고 원시 출력 [0.70,0.45,0.25]를 보정합니다.

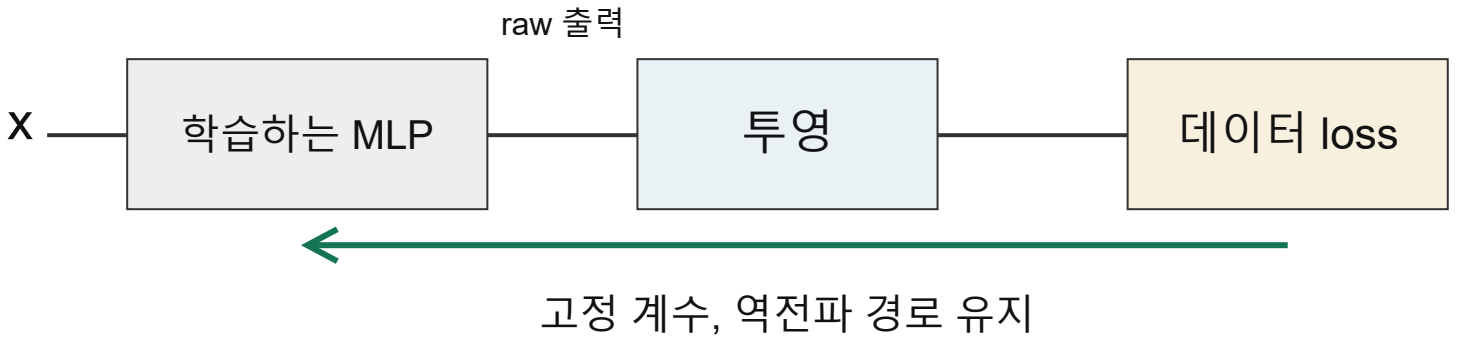
단계	계산 결과
두 잔차 $B\hat{y}-d$	[0.20,0.05]
$G=BB^{T}$의 대각 원소	첫 행 제곱합 3, 둘째 행 제곱합 2
G의 비대각 원소	1×1+1×(−1)+1×0=0
$G\lambda$=잔차	λ=[0.20/3,0.05/2]
$B^{T}\lambda$	[0.091667,0.041667,0.066667]
보정 출력	[0.608333,0.408333,0.183333]

합은 1.2이고 A−B=0.2입니다. 두 등식이 독립이어서 두 종류의 잔차를 따로 제거합니다. 두 번째 행이 첫 행의 두 배이면 G의 역행렬이 없어 같은 계산을 할 수 없습니다.

여기서 B는 제약 행렬의 이름이며 제품 성분 B와 구별합니다. 추가한 A−B 등식은 이 선형 계산 예제의 조건입니다.

학습 후 보정과 투영층을 포함한 학습

방법	학습 loss에 들어가는 출력
MLP + 학습 후 투영	Raw 예측으로 학습하고 test에서 투영합니다.
KKT-hPINN 방식	학습할 때부터 투영 출력을 loss에 넣습니다.



두 방법은 같은 등식을 만족시킬 수 있지만, gradient가 달라 학습된 파라미터는 일반적으로 다릅니다.

사후 투영은 학습이 끝난 출력의 합을 고칩니다. 학습 중 투영에서는 손실이 이미 보정된 출력을 보므로, 원시 출력 전체를 같은 만큼 늘리는 방향처럼 투영에서 지워질 방향은 출력 손실의 기울기에 기여하지 않습니다.

고정 투영층의 역전파

$$\frac{\partial y_p}{\partial \hat{y}} = P$$

$$\nabla_{\hat{y}} L = P^T \nabla_{y_p} L$$

P는 대칭인 orthogonal projector이므로 $P^T = P$입니다. 투영 행렬 P는 고정되어 있으며, 출력에서 전달된 gradient에 P^T를 곱해 raw 출력으로 전달합니다.

출력을 detach하거나 투영 계산을 no_grad로 감싸면 필요한 경로가 끊깁니다. θ에 대한 gradient는 MLP Jacobian까지 chain rule로 이어집니다.

역전파를 세 성분으로 확인하기

$$P = I - \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$P^T \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{-1}{3} \\ \frac{-1}{3} \end{bmatrix}$$

투영 출력에서 올라온 gradient가 [1, 0, 0]이면 raw 출력에 전달되는 값은 [2/3, −1/3, −1/3]입니다. 합이 0이므로 등식의 접공간 방향입니다.

실습은 두 개의 독립 등식에서도 analytic Jacobian과 double-precision gradcheck를 확인

기울기 [1,0,0]의 평균은 1/3입니다. 각 성분에서 이 평균을 빼면 [2/3,−1/3,−1/3]이 됩니다. 세 성분을 같은 만큼 올리는 방향은 합을 바꾸므로 투영층이 제거하고, 합을 유지하는 성분 간 조정 방향만 남깁니다.

출력 표준화에 따른 제약 행렬 변환

$$y = \mu + Dz, BDz = d - B\mu$$
$$D = \mathrm{diag}\left(s_1, \ldots, s_n\right)$$

μ와 양의 scale s_j는 train 데이터로만 계산합니다. 표준화 좌표 z에서 투영하려면 B 대신 BD, 우변 대신 d − Bμ를 사용해야 합니다.

표준화 공간에서 투영하면 물리 공간에서 성분별 제곱 보정량에 $1/s_j^2$의 가중치를 적용한 거리를 최소화합니다. 보정에 사용할 metric과 scale을 명시합니다.

투영 거리와 성분별 수정량

scale=[0.2,0.5,0.1]인 표준화 좌표에서 거리를 최소화하면 물리 단위 수정량은 scale의 제곱에 비례합니다. 총 잔차 0.20을 0.04:0.25:0.01로 나눕니다.

성분	scale² / 합	빼는 농도	보정 결과
A	0.04/0.30	0.026667	0.673333
B	0.25/0.30	0.166667	0.283333
C	0.01/0.30	0.006667	0.243333

```
scales = np.array([0.2, 0.5, 0.1])
raw = np.array([0.70, 0.45, 0.25])
weights = scales**2 / np.sum(scales**2)
projected = raw - (raw.sum() - 1.2) * weights
```

B는 표준편차가 커 같은 농도 수정이 표준화 거리에서 덜 비쌉니다. 따라서 더 많이 바뀝니다. 합 1.20만 확인하면 두 거리 정의의 차이가 보이지 않습니다.

수정량을 Δ_j라 두면 최소화할 거리는 $\tfrac{1}{2}\sum \Delta_j^2/s_j^2$이고 $\sum \Delta_j = -0.20$입니다. 정지조건 $\Delta_j/s_j^2+\lambda=0$에서 $\Delta_j=-\lambda s_j^2$를 얻고, 합 조건에 넣으면 $\lambda=0.20/\sum s_j^2$입니다. 따라서 각 수정량이 s_j^2에 비례합니다.

물리 단위 복원과 투영 순서

MLP는 표준화 출력 z를 만들고, μ + Dz로 mol/L를 복원한 뒤 물리 공간에서 투영합니다. 데이터 loss는 모든 성분에 공통인 $c_* = 1$ mol/L로 나눕니다.

$$\nabla_z L = D^T P^T \nabla_{y_p} L$$

D는 대각행렬이라 $D^T = D$입니다. 성분별 표준편차로 loss까지 나누는 경우에는 그 가중치도 gradient와 오차 해석에 반영해야 합니다.

PyTorch의 최소 구현

raw 단위: mol/L. feed shape: (batch, 1)

```
raw = y_mean + y_scale * net(x_scaled)
residual = raw.sum(dim=-1, keepdim=True) - feed
y_projected = raw - residual / 3.0
loss = ((y_projected - y_true) / c_scale).square().mean()
optimizer.zero_grad()
loss.backward()
optimizer.step()
```

투영에는 학습 파라미터가 추가되지 않습니다.

keepdim=True가 잔차를 (batch,1)로 남겨 각 행의 세 성분에 같은 보정량을 빼게 합니다. batch 전체의 합을 하나로 계산하면 서로 다른 운전 조건끼리 보정량을 공유하는 오류가 됩니다.

농도 합 제약과 비음수성

Raw [0.02, 0.90, 0.50] mol/L의 합은 1.42입니다. 1.20으로 투영하면 [−0.053333, 0.826667, 0.426667] mol/L가 됩니다.

합은 맞고 A는 음수. 음수만 0으로 자르면 합이 약 1.253333으로 바뀝니다.

비음수성과 합을 함께 요구하려면 simplex 투영처럼 해당 부등식까지 포함한 방법이 필요합니다. 활성 부등식이 바뀌면 gradient도 구간별로 달라집니다.

음수 성분을 포함한 투영의 재계산

원시 [0.02,0.90,0.50]에서 합 잔차 0.22를 균등하게 빼면 A가 음수가 됩니다. 합과 비음수 조건을 함께 요구하면 A=0인 경계에서 다시 계산합니다.

단계	계산
A를 0으로 고정	B+C=1.20
남은 원시 합	0.90+0.50=1.40
두 성분의 보정	(1.40−1.20)/2=0.10
결과	[0,0.80,0.40] mol/L

두 양의 성분에는 같은 0.10을 빼고, A는 0에서 멈춥니다. 이 결과는 원시 출력에 대한 simplex 투영입니다. 먼저 음수만 0으로 자르는 방식은 B와 C를 재조정하지 않아 합을 맞추지 못합니다.

원시 값이 변해 A가 다시 양수가 되는 경계에서는 활성 제약이 바뀝니다. 따라서 이 변환의 미분은 고정 선형 등식 투영처럼 전 구간에서 같은 행렬이 아닙니다.

투영의 오차 감소 조건

기준 y^*가 같은 affine 집합에 있으면 Euclidean 투영의 직교성으로:

$$\left\|\hat{y} - y^*\right\|^2 = \left\|\hat{y} - y_p\right\|^2 + \left\|y_p - y^*\right\|^2$$

따라서 같은 raw 예측을 물리 공간에서 투영하면 전체 제곱오차는 증가하지 않습니다. 성분 하나의 오차는 커질 수 있습니다.

측정 잡음으로 기준값이 집합 밖에 있거나 다른 가중 metric을 사용하면 이 물리 Euclidean 비교를 그대로 적용할 수 없습니다. 별도로 학습한 두 모델의 우열도 이 식에서 나오지 않습니다.

원시 오차는 제약면에 수직인 합 오차와 제약면 안의 성분 분배 오차로 나뉩니다. 투영은 첫 부분만 제거합니다. 직교하는 두 부분의 제곱합이 원래 오차이므로, 같은 기준값과 같은 거리에서 전체 제곱 오차가 늘지 않습니다.

비교 실험의 조건을 고정하기

입력 영역은 $T = 305\text{-}355$ K, $\tau = 1\text{-}8$ min, $C_{Af} = 0.8\text{-}1.5$ mol/L. 속도식은 $k_1 = 0.2 \exp[6000(1/330 - 1/T)]$, $k_2 = 0.1 \exp[5000(1/330 - 1/T)]$ min^{-1}

Train/validation/test = 900/300/300행, seed = 404, 잡음 없는 analytic label을 사용합니다. 모델은 모두 3 → 32 → 32 → 3 ReLU MLP(1283개 파라미터)입니다.

초기 가중치·batch 순서·train 통계량을 공유합니다. Adam 0.001, batch 128, 600 epoch 중 validation 물리 MSE가 최소인 가중치를 선택하고 test를 평가

실행 결과: 정확도와 등식 위반

모델	전체 RMSE	최대 합 위반	음수 항목 수
MLP	0.003657	1.17e-02	1
Soft (λ = 10)	0.008161	6.30e-03	5
MLP + post	0.003426	2.22e-16	1
Hard	0.003036	3.33e-16	1

농도·오차 단위: mol/L. 음수 집계 기준: 900개 test 성분값 중 -10^{-8} 미만

이번 실행에서 soft penalty는 합 위반을 줄였지만 RMSE가 커졌습니다. Hard와 post는 합을 맞췄고 음수 농도는 남았습니다. 한 seed의 결과이므로 일반적인 정확도 순위를 뜻하지 않습니다.

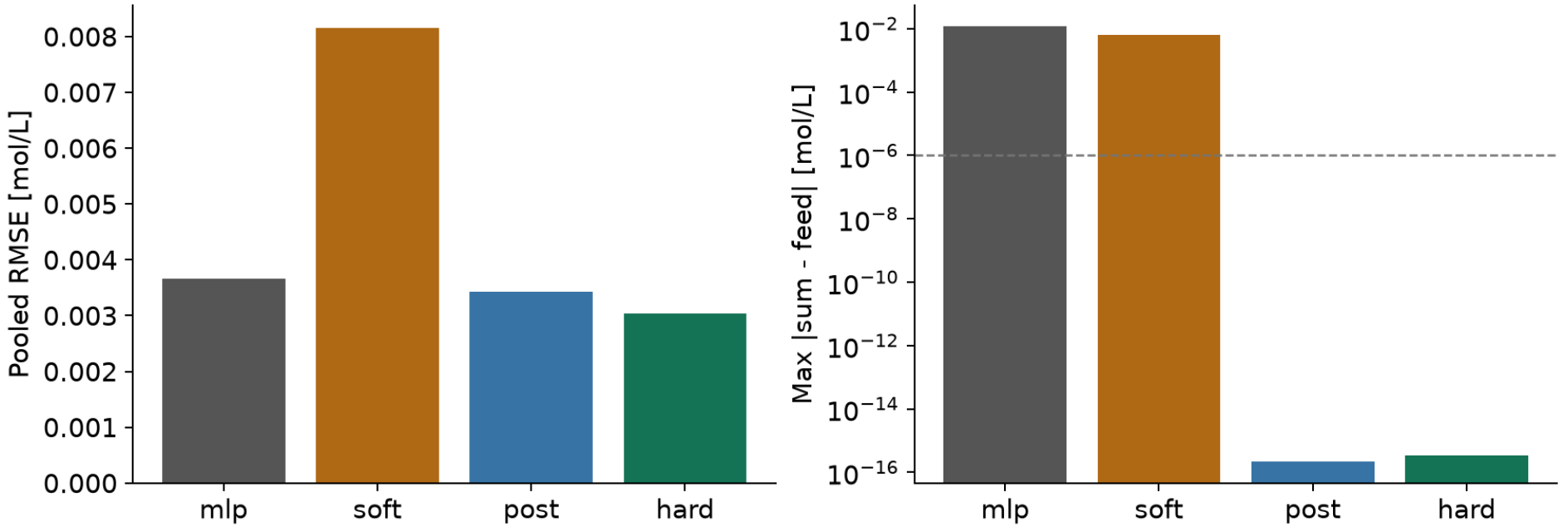
성분별 RMSE와 최소 농도

모델	A RMSE	B RMSE	C RMSE	최소 농도
MLP	0.003866	0.003383	0.003704	-0.003451
Soft (λ = 10)	0.008753	0.008742	0.006840	-0.010393
MLP + post	0.003724	0.003015	0.003500	-0.004176
Hard	0.003728	0.002540	0.002703	-0.002822

단위는 모두 mol/L입니다. 전체 RMSE만 보면 성분별 차이를 놓칠 수 있습니다.

최소 농도는 세 성분과 test 행 전체에서 구합니다. 평가 tolerance와 실제 최소값을 함께 기록

학습에 포함하지 않은 성분수지의 평가



Hard 모델은 농도 합 등식을 만족하며, 성분별 반응·유출 관계에는 residual RMS 0.002225 mol/(L·min)가 남았습니다.

합 잔차 0은 세 종별 잔차의 합이 0이라는 조건과 연결됩니다. 양수와 음수 종별 잔차가 서로 지워질 수 있으므로 종별 RMS는 0이 아닐 수 있습니다. 그림은 학습에 넣지 않은 이 추가 조건을 검사합니다.

실습 실행과 재현 기록

실습 ZIP, Notebook, Python 스크립트를 제공합니다. ZIP을 풀고 같은 폴더에서 실행합니다.

```
python -m pip install -r requirements.txt
python week04_lab.py --check-only
python week04_lab.py --output-dir week04-results
```

결과에는 분할된 dataset, test 예측, 모델 가중치, train 통계량, validation 이력과 평가 JSON이 저장됩니다. 제공 수치는 CPU·float64·PyTorch 2.14.1+cpu에서 실행했습니다.

동적 모델의 총 농도 등식

구간 안에서 유입 농도와 τ가 일정하면 총 농도의 정확한 전이는:

$$S_{k+1} = C_{\mathrm{Af}} + \left(S_k - C_{\mathrm{Af}}\right) e^{-\frac{\Delta t}{\tau}}$$

초기 합 S_k가 C_Af와 다르면 다음 합을 곧바로 C_Af로 강제하면 안 됩니다. 다음 상태의 B = [1, 1, 1]은 같지만 우변은 위 전이식입니다.

τ를 입력으로 두면 등식의 우변 d(x)는 τ에 비선형으로 의존합니다. 이 우변을 계산한 뒤 고정 B를 사용하는 투영을 적용합니다.

연습문제와 제출물

1. [0.70, 0.45, 0.25]를 합 1.20에 투영하고 gradient [1, 0, 0]의 역전파를 계산
2. B의 두 행이 [1, 1, 1], [2, 2, 2]일 때 d = [1.2, 2.5]의 가능 여부를 설명합니다.
3. λ = 0.1, 1, 10을 validation으로 비교하고, 선택한 설정만 test에서 평가
4. 각 모델의 A/B/C RMSE·합 MAE/최대값·최소 농도·성분수지 residual을 제출

다음 주: 학습한 ReLU predictor의 고정과 최적화 문제 정의

읽기자료와 논문

4주차 읽기자료 · 국문 읽기 PDF

Chen, H., Constante Flores, G. E., & Li, C. (2024). *Physics-informed neural networks with hard linear equality constraints*. Computers & Chemical Engineering, 189, 108764. DOI · 공개 원고 · 저자 코드·데이터.

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). *Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations*. Journal of Computational Physics, 378, 686-707. DOI.

수학·구현 원문

Boyd, S., & Vandenberghe, L. (2004). *Convex Optimization*. 5.5.3절의 KKT 조건, 10.1절의 등식 제약 최소화. 공식 교재.

PyTorch 공식 문서: 선형계 풀이, 수치 gradient 확인. Raissi et al., 2017 Part I 공개 원고.