7-Дәріс

Өнімділік беттері және оңтайлы нүктелер

Дәрістің мақсаты

Студенттерде жасанды нейрондық желілерді оқыту процесін математикалық оңтайландыру міндеті ретінде қарастыру дағдысын қалыптастыру, өнімділік (қателік) беттері ұғымын, көпөлшемді кеңістіктегі Тейлор қатарын, градиент пен Гессиан матрицасын, сондай-ақ локальды және глобальды минимумдарды анықтаудың бірінші және екінші ретті шарттарын меңгерту.

Оқу нәтижелері

Дәрісті меңгеру нәтижесінде студенттер:

- **Түсінеді:** Нейрондық желінің өнімділік функциясы (қателік беті) ұғымын, стационарлық нүктелердің түрлерін (локальды минимум, глобальды минимум, ер-тоқым нүктесі) және Тейлор қатарының көпөлшемді векторлық түрін.
- **Есептей алады:** Берілген мақсатты функцияның градиент векторын ($\nabla F(x)$) және екінші ретті дербес туындылар матрицасын — Гессианды ($\nabla^2 F(x)$).
- **Талдай алады:** Бірінші ретті ($\nabla F(x^*) = 0$) және екінші ретті (Гессиан матрицасының оң анықталғандығы) оңтайлылық шарттарын қолданып, табылған нүктенің қатаң локальды минимум екенін математикалық түрде дәлелдей алады.

1. Өнімділік беттері (Performance Surfaces) ұғымы

- **Оңтайландыру міндеті ретінде оқыту:** Нейрондық желіні оқытудың негізгі мақсаты — желінің қателігін (немесе шығын функциясын) барынша азайтатын оңтайлы салмақтар мен ығысулар векторын (x) табу.
- **Өнімділік беті:** Желі параметрлерінің (x) мәндеріне тәуелді өнімділік көрсеткішінің ($F(x)$) графигі көпөлшемді кеңістіктегі **өнімділік бетін** құрайды. Оқыту процесі — осы беттің ең төменгі нүктесін (минимумын) іздеу.
- **Градиент векторы ($\nabla F(x)$):** Көп айнымалылы $F(x) = F(x_1, x_2, \ldots, x_n)$ функциясының бірінші ретті дербес туындыларынан тұратын вектор:

$$\nabla F(x) = \left[\frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2}, \ldots, \frac{\partial F}{\partial x_n} \right]^T$$

Геометриялық тұрғыда градиент функцияның **ең жылдам өсу** бағытын көрсетеді, ал антиградиент ($-\nabla F(x)$) — **ең жылдам кему** бағытын көрсетеді.

2. Векторлық жағдай және көпөлшемді Тейлор қатары

- **Градиент векторы ($\nabla F(x)$):** Көп айнымалылы $F(x) = F(x_1, x_2, \ldots, x_n)$ функциясының бірінші ретті дербес туындыларынан тұратын вектор:

$$\nabla F(x) = \left[\frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2}, \ldots, \frac{\partial F}{\partial x_n} \right]^T$$

- Егер x^* нүктесі $F(x)$ дифференциалданатын функциясының локальды минимумы болса, онда осы нүктеде функцияның бірінші ретті туындысы (градиенті) міндетті түрде нөлге тең болуы керек:

$$\nabla F(x^*) = 0$$

- Бұл шартты қанағаттандыратын нүктелер **стационарлық (критикалық) нүктелер** деп аталады. Алайда бұл шарт жеткіліксіз, өйткені максимум мен ер-тоқым нүктелерінде де градиент нөлге тең болады.

- Бұл шартты қанағаттандыратын нүктелер **стационарлық (критикалық) нүктелер** деп аталады. Алайда бұл шарт жеткіліксіз, өйткені максимум мен ер-тоқым нүктелерінде де градиент нөлге тең болады.

5. Екінші ретті оңтайлылық шарттары (жеткілікті шарт)

- Стационарлық x^* нүктесінің ($\nabla F(x^*) = 0$) қатаң локальды минимум екеніне көз жеткізу үшін Гессиан матрицасының ($\nabla^2 F(x^*)$) **оң анықталған (positive definite)** болуы тексеріледі:

$$\Delta x^T \nabla^2 F(x^*) \Delta x > 0 \quad (\text{кез келген } \Delta x \neq 0 \text{ үшін})$$

- **Меншікті мәндер критерийі:** Симметриялы Гессиан матрицасы оң анықталған болады сонда және тек сонда ғана, егер оның **барлық меншікті мәндері қатаң оң болса** ($\lambda_i > 0$). Егер меншікті мәндерінің таңбалары әртүрлі болса (бірі оң, бірі теріс), онда x^* — ер-тоқым нүктесі.

Қорытынды

Бұл дәрісте нейрондық желілерді оқытудың математикалық іргетасы болып табылатын өнімділік беттері мен көпөлшемді оңтайландыру теориясы қарастырылды. Тейлор қатарының векторлық моделі арқылы функцияның локальды мінез-құлқын талдауда градиент векторы мен Гессиан матрицасының шешуші рөлі көрсетілді.

Бірінші ретті (қажетті) және екінші ретті (жеткілікті) оңтайлылық шарттарын меңгеру стационарлық нүктелерді дұрыс жіктеуге (локальды минимумды ер-тоқым нүктесінен ажыратуға) мүмкіндік береді. Бұл теориялық білім алдағы дәрістерде қарастырылатын градиенттік түсу, Уидроу-Хофф (LMS) және қатені кері тарату (Backpropagation) алгоритмдерінің жұмыс принципін түсіну үшін аса қажет.

Өзін-өзі бағалау сұрақтары

1. **Өнімділік беті:** Нейрондық желілер теориясында «өнімділік беті» дегеніміз не және оқыту процесі геометриялық тұрғыда қалай сипатталады?

2. **Градиент және Гессиан:** Көп айнымалылы (F(x)) функциясының *градиент векторы* мен *Гессиан матрицасы* қалай анықталады? Гессиан матрицасы өнімділік бетінің қандай қасиетін сипаттайды?

3. **Бағытталған туынды:** Егер біз (x) нүктесінен бастап функция мәнін ең жылдам азайтқымыз келсе, қай бағытта қозғалуымыз керек? Жауабыңызды математикалық түрде негіздеңіз.

4. **Оңтайлылық шарттары:** Стационарлық нүктенің *қатаң локальды минимум* болуы үшін қандай бірінші ретті және екінші ретті шарттар орындалуы тиіс?

5. **Меншікті мәндер рөлі:** Гессиан матрицасының меншікті мәндері бойынша табылған стационарлық нүктенің *минимум*, *максимум* немесе *ер-тоқым нүктесі* екенін қалай анықтауға болады?

Әдебиеттер тізімі

1. **Хейган М. Т., Демут Х. Б., Бил М. Х., Де Хесус О.** Нейронные сети. Дизайн и применение (Neural Network Design). — М.: ДМК Пресс, 2021. — 800 с. *(8-тарау: «Поверхности показателей качества и точки оптимума» — векторлық жағдай, Тейлор қатары, бірінші және екінші ретті оңтайлылық шарттары).*

2. **Гудфеллоу Я., Бенджио И., Курвилль А.** Глубокое обучение. — М.: ДМК Пресс, 2018. — 652 с. *(4-тарау: «Численные вычисления» — градиенттік оңтайландыру, Гессиан матрицасы және ер-тоқым нүктелері).*

3. **Хайкин С.** Нейронные сети: полный курс. — 2-е изд. — М.: Издательский дом «Вильямс», 2018. — 1104 с. *(Оңтайландыру әдістері және қателік беттерін талдау).*

4. **Васильев Ф. П.** Методы оптимизации. — М.: МЦНМО, 2011. — 620 с. *(Көпөлшемді функциялардың экстремум шарттары және оң анықталған квадраттық формалар).*

5. **Бейсембекова Р. Н., Әбдірахманова Ж. Б.** Жасанды интеллект және нейрондық желілер негіздері: Оқу құралы. — Алматы: Қазақ университеті, 2021. — 184 б. *(Қазақ тіліндегі оңтайландыру негіздері мен градиенттік әдістер).*