



EÖTVÖS LORÁND TUDOMÁNYEGYETEM
INFORMATIKAI KAR
NUMERIKUS ANALÍZIS TANSZÉK

Adaptív transzformációs módszerek a jel- és képfeldolgozásban

HABILITÁCIÓS TÉZISFÜZET

Dr. KOVÁCS Péter
kovika@inf.elte.hu

Budapest, 2021

Tartalomjegyzék

Előszó	ii
1. Bevezetés	1
1.1. Jelek matematikai reprezentációja	2
2. Jelmodellezés	4
2.1. Adaptív projekciós módszerek	4
2.1.1. Spline-ok illesztése az alappontok variálásával	5
2.1.2. Paraméterezett Hermite-függvények	6
2.1.3. Racionális függvények pólusoptimalizációja	8
2.2. Ritka reprezentációk	9
3. Gépi tanulás	10
3.1. Hibrid tanuló algoritmusok	12
3.1.1. Termografikus képrekonstrukció	13
3.2. Modell-vezérelt neurális hálók	15
4. Alkalmazások	17
4.1. EKG, EEG jelfeldolgozás	17
4.2. Rendszeridentifikáció	20

Előszó

Tudományos pályámat 2010-ben kezdtem PhD hallgatóként az ELTE IK Numerikus Analízis Tanszékén, ahol 2012-től tanársegéd, majd 2016-tól adjunktusként dolgozom. Ez idő alatt bekapcsolódhattam a tanszéken működő, több évtizedes múltra visszatekintő Harmonikus- és Fourier-analízis Iskola kutatásaiba. Dr. Schipp Ferenc professzor emeritus és Dr. Fridli Sándor tanszékvezető egyetemi tanár témavezetése nyomán megismertem a racionális függvényrendszerekkel való approximáció elméleti eredményeket, ezek geometriai értelmezését, rendszer- és irányításelméleti jelentőségét, illetve EKG jelfeldolgozásban történő gyakorlati alkalmazását. Ez az időszak tudományos és oktatói szempontból is meghatározó volt. Ennek köszönhetően doktori tanulmányaim alatt és későbbi tudományos pályám során is az említett témakörökhöz erősen kapcsolódó alkalmazásorientált alapkutatásokat végeztem, melyekben az elméleti kutatás problémafelvetéseit közvetlen gyakorlati alkalmazások inspirálták. Fontosnak tartom kiemelni, hogy a tanszéken oktatott tárgyak nagy mértékben hozzájárultak kutatásaim sikerességéhez, és fordítva, bizonyos eredményeim, észrevételeim segítettek ezen tárgyak ismeretanyagának mélyebb megértésében, melyet oktatási segédanyagok, szemléltető programok formájában beépítettem egyetemi előadásaimba. Az oktatott tárgyak anyagaira támaszkodó, saját kutatásaimhoz kapcsolódó feladatokon keresztül egyetemi hallgatókat is be tudtam vonni a tudományos munkába. Ennek köszönhetően az elmúlt években összesen 4 diplomamunkát, 10 szakdolgozatot és 4 TDK dolgozatot témavezettem, melyek az OTDK-n is szépen szerepeltek (II. és III. díj, illetve egy különdíj). 2020-ban az egyik ilyen TDK dolgozat folytatásaként egy doktori témát is meghirdettem.

Az elmúlt időszakban nemzetközi tapasztalatokat is szereztem. 2012-ben fél évet töltöttem Finnországban, a Tamperei Műszaki Egyetem Jelfeldolgozó Kutató Intézetében (ISP). Itt epilepsziás rohamok EEG alapú detektálásán dolgoztam. A kutatócsoportot vezető Moncef Gabbouj professzorral és PhD hallgatójával Kaveh Samiee-val több közös publikációnk jelent meg, köztük az egyik egy D1-es folyóiratban [SamKov17]. Finnországban ismertem meg Dr. Serkan Kiranyaz-t a Katar-i Egyetem professzorát is, akivel sztochasztikus optimalizáció témakörében jelent meg közös cikkünk. Ez és az ELTE Numerikus Analízis Tanszékén végzett hiperbolikus geometriával kapcsolatos kutatások inspirálták egy későbbi munkám, az adaptív transzformációs módszerek optimalizációjának általánosítását, mely az egyik legrangosabb jelfeldolgozó folyóiratban jelent meg D1-es publikációként [KovFri20]. A cikk eredményeit bemutattam Victor Pereyra és Michael Saunders stanfordi

professzoroknak is, akik 2017-ben meghívtak a *Linear Algebra and Optimization* szemináriumukra.

2018 és 2020 között a linzi Johannes Kepler Egyetem (JKU) Jelfeldolgozó Intézetében voltam posztdoktori kutató. Itt többek között folytathattam a doktori alatt megkezdett EKG jelfeldolgozással kapcsolatos kutatásaimat. Ezen keresztül ismertem meg Carl Böck PhD hallgatót és a Kepler Egyetemi Kórház Intenzív és Aneszteziológiai Osztályának vezetőjét, Dr. Jens Meier orvosprofesszort, akikkel EKG jelek modellezését és szegmentálását kutattuk. Az EKG és EEG jelfeldolgozáshoz kapcsolódóan 2019-ben egy közös, ELTE IK – JKU ISP, konferencia szekciót is szerveztük a 17. *EUROCAST* konferencián. Jelenleg egy hasonló témában szervezünk szekciót az 55. *Asilomar Conference on Signals, Systems, and Computers* konferenciára. A biológiai jelfeldolgozás mellett, korábbi kutatásaimtól eltérő jelfeldolgozási problémákon (telekommunikáció, gépi tanulás, roncsolásmentes anyagvizsgálat) is dolgozhattam a linzi intézetben. Ennek eredményeképp az ISP-t vezető Dr. Mario Huemer professzorral 9 közös publikációnk jelent meg köztük kettő Q1-es folyóiratokban. Egy egyetemi együttműködés keretében a 2019-ben újonnan megalakult Silicon Austria Labs GmbH indulóprojektjeinek kidolgozásában is részt vehettem. Így a kutatás mellett számos új, értékes tapasztalat gazdagodtam a projektvezetés, szervezés, pályázatírás területén.

Doktori védésem óta, a 2016-2021 időszakban 6 konferencia és 11 folyóirat cikket publikáltam jelmodellezés, mesterséges intelligencia/gépi tanulás, biológiai jelfeldolgozás, telekommunikáció, roncsolásmentes anyagvizsgálat témakörökben. Ezek alapvetően interdiszciplináris jellegű komplex kutatások, melyekhez nem csak informatikai, illetve matematikai tudásra, de alkalmazásspecifikus ismeretekre is szükség van. Emiatt különböző területeken dolgozó szakemberekkel, orvosokkal, mérnökökkel, fizikusokkal kellett együttműködnöm. Habilitációs dolgozatomban ezen időszak kutatási eredményeit foglalom össze, rendszerezem három témában:

- jelfeldolgozásban: [Kov17b; Kov18; KovFek19; KovBöc19; BogFri19; KovFri20];
- gépi tanulásban: [KovLeh20a; KovLeh20b; LehGal21; KovBog21b];
- gyakorlati alkalmazásokban: [KovBöc17; SamKov17; LanKov18; BöcKos19; DózBog19; GarKov20; BöcKov21].

A habilitációs dolgozat szerkesztése inkrementális, tehát építék a korábban bevezetett fogalmakra és állításokra, éppen ezért a fejezetek sorrendisége fontos. Ezt a logikát követi a publikációk csoportosítása, melyben figyelembe vettem az alkalmazott matematikai módszerek, technikák hasonlóságát is.

Köszönetnyilvánítás

Ezúton szeretnék köszönetet mondani Dr. Fridli Sándor tanszékvezető egyetemi tanárnak, hogy kezdettől fogva koordinálta tudományos pályafutásomat, témavezette doktori tanulmányaimat, és támogatta habilitációs dolgozatom elkészítését is. Hálás vagyok a közös publikációk alatt szerzett tapasztalatokért. Köszönöm a tudományos ösztöndíjak és pályázatok megírásához nyújtott segítségét és szakmai támogatását.

Külön köszönettel tartozom Dr. Schipp Ferenc professzor emeritusznak, aki a doktori tanulmányaim megkezdése óta mentorálja tudományos pályám és ötleteivel azóta inspirálja immáron nemcsak saját, de doktoranduszom kutatásait is.

Köszönöm az ELTE Numerikus Analízis Tanszékének, hogy 2012-ben támogatva tanársegédi, 2016-tól adjunktusi kinevezésem, illetve jelenlegi habilitációs pályázatomat. Hálás vagyok a tanszéki kollégáknak, hogy a 2018-2020 közötti osztrák kiküldetésem alatt helyettem is viselték a Tanszék oktatási terheit és ezáltal lehetővé tették, hogy a külföldi posztdoktori időszak után visszatérhessek az ELTE-re.

Hálás vagyok Dr. Kaveh Samiee-nek és Carl Böck-nek az elmúlt évek közös munkájáért és barátságukért. Kaveh-val a Tamperei Műszaki Egyetemen találkoztam. PhD hallgatóként ez volt az első nemzetközi együttműködésem, melyből a jelenleg legtöbbet hivatkozott publikációim készültek. Carl-t, a linzi Johannes Kepler Egyetem (JKU) PhD hallgatóját, már posztdoktori kutatóként ismertem meg. Ez tudományos pályám egyik mérföldköve, melynek köszönhetően posztdoktori állást kaptam a JKU-n. A közös publikációk, konferencia szereplések, tudományos utazások, mellett Carl jóvoltából végül síelni is megtanultam.

Köszönöm családomnak a doktori tanulmányaim és a külföldi kiküldetésem során nyújtott támogatását. A dolgozatot édesanyám emlékének ajánlom.

1. fejezet

Bevezetés

Az informatika tudományos és technológiai vívmányainak köszönhetően a jelfeldolgozás óriási fejlődésen ment keresztül az elmúlt több mint 50 évben [Neb98]. A kezdetben kizárólag hadászati és telekommunikációs célokat szolgáló jelfeldolgozás mára számos további tudományos és köznapi alkalmazással (szeizmológia, biológiai jel- és képfeldolgozás, idősor analízis, beszédfelismerés, stb.) bővült. Ezekben az esetekben a vizsgált jelet független változók (pl. idő, tér) függvényének tekinthetjük. A dolgozatban tárgyalt jelek többsége egyváltozós valós-valós függvény, de

például a roncsolásmentes anyagvizsgálathoz kapcsolódó munkáimban előforduló kép és videó felvételeket két-, illetve háromváltozós függvényekkel reprezentálhatjuk. A gyakorlatban előforduló analóg jelek folytonosak, melyeket a számítógépes feldolgozhatóság érdekében digitalizálni kell, ami magában foglalja az értelmezési tartomány mintavételezését és az értékkészlet kvantálását. Az analóg-digitális átalakítás eredményeképp függvények helyett vektorokkal is dolgozhatunk. Ha a mintavételezést megfelelő gyakorisággal végezzük, akkor a Shannon–Nyquist-féle mintavételezési tétel értelmében a diszkrét mérésekből hibátlanul rekonstruálhatjuk az eredeti jelet. Tehát nem veszítünk információt, ha az analóg jelek helyett azok diszkrét megfelelőjével dolgozunk.

Attól függően, hogy a mérésekben feltételezünk-e valamilyen bizonytalanságot, beszélhetünk determinisztikus és nemdeterminisztikus jelekről. Az első esetben a jel akár explicit képlettel is definiálható, míg a második esetben a méréseket befolyásoló véletlen tényezők (pl. zaj) miatt erre nincs lehetőség. A nemdeterminisztikus jelek sztochasztikus folyamatoknak tekinthetők, a diszkrét mérések pedig valószínűségi vektorváltozókkal modellezhetők. Így a jelfeldolgozásbeli problémák tárgyalása alapvetően két irányból, approximáció- és valószínűségelméleti alapon közelíthető meg. Az ELTE IK képzéseiben az előbbi szemlélet dominál köszönhetően a Numerikus Analízis Tanszéken folyó több évtizedes approximációelméleti kutatásoknak. Ezért a habilitációs tézisekben összefoglalt saját eredményeimet ebben a megközelítésben fogom ismertetni. Megjegyzem, hogy a JKU-n volt szerencsém a valószínűségelméleti alapú jelfeldolgozást is oktatni az "*Optimum and Adaptive Signal Processing Systems*" című egy féléves MSc előadáson keresztül.

1.1. Jelek matematikai reprezentációja

A jelek feldolgozásához szükség van egy általános keretrendszerre amiben az általuk hordozott információt értelmezni tudjuk. Ehhez először kijelölünk egy, a problémához illeszkedő $(\mathcal{H}, \|\cdot\|)$ normált teret. A számítógépes feldolgozáshoz ez önmagában nem elég, ezért az analitikusan nehezen kezelhető bonyolult $f \in \mathcal{H}$ függvényeket egyszerűbb leképezésekkel közelítjük. Formálisan a következő feladat megoldásáról van szó:

$$(1.1) \quad d(f, \mathcal{S}) := \inf_{g \in \mathcal{S}} \rho(f, g) = \inf_{g \in \mathcal{S}} \|f - g\| \quad (f \in \mathcal{H}).$$

ahol $\mathcal{S} \subset \mathcal{H}$ az alkalmazáshoz választott altér. A fenti $\rho(f, \cdot)$ távolságot minimalizáló $\tilde{f} \in \mathcal{S}$ elem létezésére, egyértelműségére, karakterizálására számos klasszikus approximációelméleti tétel ismert [Szi07; Nat52]. A gyakorlatban az $f \in \mathcal{H}$ függvény egy véges, számítógépes tárolásra alkalmas $\tilde{f} \in \mathcal{S} \subset \mathcal{H}$ reprezentációjára van szükségünk. Ezért a legjobb approximáció problémakörének egy speciális esetére szorítkozunk, melyben \mathcal{H} egy Hilbert-tér a $\langle \cdot, \cdot \rangle$ skaláris szorzattal, $\|f\| = \sqrt{\langle f, f \rangle}$, és \mathcal{S} egy véges dimenziós altere \mathcal{H} -nak. Ez többek között azért előnyös, mert a létezés és egyértelműség mellett biztosított az f -et legjobban közelítő \tilde{f} elem elő-

állítás is:

$$(1.2) \quad \tilde{f}(t) = P_{\Theta}^N f(t) := \sum_{j=0}^{N-1} c_j \Theta_j(t) \quad (t \in \mathbb{R}),$$

ahol P_{Θ}^N az $\{\Theta_j : 0 \leq j < N\}$ függvények által kifeszített \mathcal{S} altérre való merőleges vetítést leíró operátor. A $c_j \in \mathbb{R}$ ($j = 0, \dots, N-1$) együtthatókat a $\mathbf{G}\mathbf{c} = \mathbf{b}$ lineáris egyenletrendszer megoldásaként kapjuk, melyben Az egyenlet bal oldalán egy ún. Gram mátrix $\mathbf{G} = (\langle \Theta_i, \Theta_j \rangle)_{i,j=0}^{N-1}$ áll, a jobb oldali vektor pedig $\mathbf{b} = (\langle f, \Theta_0 \rangle, \dots, \langle f, \Theta_{N-1} \rangle)^T$ alakú. Ha $\{\Theta_j : 0 \leq j < N\}$ egy ortonormált függvényrendszer, akkor \mathbf{G} az egységmátrix és a $c_i = \langle f, \Theta_i \rangle$ számokat az f függvény Fourier-együtthatóinak nevezzük. Ez már egy számítógépes tárolása alkalmas reprezentációja az f jelnek, melynek így elég csak a c_i együtthatóival műveletet végezni (pl. tömöríteni, szűrni). Az \mathcal{S} alteret és az azt definiáló $\{\Theta_j : 0 \leq j < N\}$ függvényrendszert alkalmazástól függően választjuk meg. Például hangfeldolgozásban a trigonometrikus rendszer elemeit használjuk (MP3), képfeldolgozásban a diszkrét koszinusz, illetve wavelet transzformációk terjedtek el (JPEG, JPEG2000), telekommunikációban pedig a Walsh-transzformációnak van nagy jelentősége (CDMA). Megjegyzem, hogy a jelek (1.2)-ben vázolt értelmezése kellően általános, melyből a valószínűségelméleti megközelítések is levezethetők. Ebben az esetben a \mathcal{H} Hilbert-tér véges szórású valószínűségi változók halmaza, a skaláris szorzat pedig a $\langle x, y \rangle := E(x, y)$ alakú, ahol $E(\cdot)$ a szokásos várható érték operátort jelöli. Számos, a becsléelméletben jól ismert eljárást származtathatjuk ilyen módon [Kay93].

A dolgozat 2. fejezetében az (1.2)-hez hasonló reprezentációk nemlineáris általánosításait vizsgálom, melyekben a cél a jelek pontos ugyanakkor ritka reprezentációja. Utóbbi azt jelenti, hogy minimalizáljuk az f közelítéséhez felhasználható elemi függvények számát. Az itt kidolgozott adaptív reprezentációs módszereket a 3. fejezetben beépítem neurális hálókba is. Az így kapott modell-vezérelt neurális hálók öröklik az adaptív projekciós operátorok jó tulajdonságait, ami értelmezhető paramétereket, és kompakt háló struktúrát eredményez. Az utolsó, 4. részben az első két fejezet eredményeihez kapcsolódó gyakorlati alkalmazásokat ismertetem.

2. fejezet

Jelmodellezés

A jelek, illetve függvények hatékony reprezentációjának kutatása nem újkeletű dolog, az elmúlt évszázadban számos elméleti és gyakorlati jelentőségű eredmény született ebben a témában. Egy reprezentáció hatékonyságát két alapvető szempont, a pontosság és a komplexitás, szerint mérhetjük. E két tulajdonság fontosságát már a 14. században felismerték, amire a mai napig Ockham borotvajaként hivatkozunk. Az elv szerint, egy adott jelenséget egyformán jól leíró magyarázatok közül a legegyszerűbbet érdemes választani. Esetünkben egy f jel két egyformán valósághű \tilde{f}_1 és \tilde{f}_2 reprezentációja közül a kevesebb adattal leírhatót célszerű használni. Formálisan, a pontosságon az eredeti $f \in \mathcal{H}$ jel és hozzá az (1.1) modell szerint legközelebbi $\tilde{f} \in \mathcal{S}$ reprezentáció $\|f - \tilde{f}\|$ hibáját, míg komplexitáson az \tilde{f} előállításához szükséges információ mennyiségét értjük. Utóbbit jellemezhetjük például \tilde{f} ritkaságával, azaz $\|\mathbf{c}\|_0$ -al, ami a (1.2) egyenletben a nem nulla c_j együtthatók száma. Ez a két tulajdonság, hasonlóan sok más, a számítástudományokból jól ismert mennyiséghez, csak egymás rovására javítható. Ha például a modellezéshez használt \mathcal{H} térben teljes ortonormált $\{\Theta_j : j \in \mathbb{N}\}$ függvényrendszerek szerinti sorfejtések N -ik részletösszegeit vizsgáljuk, akkor $\lim_{N \rightarrow \infty} \|f - \tilde{f}_N\| = 0$, tehát nagy N -ekre elhanyagolhatóan kicsi hibájú reprezentációkat kaphatunk, viszont egyre több c_i együtthatót is kell tárolni. Ebben a fejezetben olyan numerikus optimalizációs eljárásokkal foglalkozom, melyek bizonyos kompromisszumok mellett egyszerre próbálják meg kielégíteni a pontosság és a komplexitás kritériumait.

2.1. Adaptív projekciós módszerek

A jelfeldolgozás első lépéseinek egyike a kérdéses probléma egyszerűsítése, felbontása kisebb részfeladatokra, melyekben már a jelnek egy egyszerűbb alakjával dolgozunk tovább. Ennek egyik lehetséges megközelítése a fizikai jelenséget leíró törvényszerűségek linearizációja, illetve a vizsgált jelek egyszerűsítése, felbontása egy előre megválasztott függvényrendszer szerint, ahol a feldolgozandó jeleket véges sok alapfüggvény lineáris kombinációjával közelítjük az (1.1)-(1.2) egyenletek szerint. A $\{\Theta_j : 0 \leq j < N\}$ rendszer rögzítése helyett azonban célszerű olyan függvényeket választani, melyek szabad paraméterekkel is rendelkeznek. Ezek az ún. adaptív

projekciós módszerek, melyekben a szabad paramétereket az alkalmazáshoz illeszkedően, optimalizációs eljárások segítségével választjuk meg. Ehhez tekintsük a $\Theta(\mathbf{a}) := \{\Theta_j(\cdot, \mathbf{a}) \in \mathcal{H} : 0 \leq j < N\}$ alakú, az $\mathbf{a} \in \Gamma \subset \mathbb{R}^n$ vektorral paraméterezett függvényrendszereket, és az (1.2) modell alábbi általánosítását:

$$(2.1) \quad \tilde{f}(t, \mathbf{a}) := P_{\Theta(\mathbf{a})}^N f(t) = \sum_{j=0}^{N-1} c_j(\mathbf{a}) \Theta_j(t, \mathbf{a}) \quad (t \in \mathbb{R}),$$

ahol a $c_j(\mathbf{a})$ együtthatókat a $\mathbf{G}(\mathbf{a})\mathbf{c}(\mathbf{a}) = \mathbf{b}(\mathbf{a})$ lineáris egyenletrendszer megoldásával kapjuk. A $\mathbf{G}(\mathbf{a})$ Gram mátrix és a jobb oldali $\mathbf{b}(\mathbf{a})$ vektor elemeit az (1.2)-ben ismertetett módon skaláris szorzatok segítségével számoljuk ki, a különbség annyi, hogy ezek most függenek az \mathbf{a} paramétertől. Így az (1.1) egyenletben definiált optimalizációs probléma a következőképpen módosul:

$$(2.2) \quad \min_{\mathbf{a}} d(f, \mathcal{S}(\mathbf{a})) := \min_{\mathbf{a}} \min_{g \in \mathcal{S}(\mathbf{a})} \|f - g\| \quad (f \in \mathcal{H}).$$

ahol $\mathcal{S}(\mathbf{a}) := \text{span}\{\Theta_j(\cdot, \mathbf{a}) : 0 \leq j < N\}$. Megjegyzem, hogy a külső minimum általában nem létezik, gyakorlati alkalmazásokban azonban az \mathbf{a} paraméterek fizikai jelentését felhasználva a keresési tartomány leszűkíthető egy kompakt halmazzá. Ezért a továbbiakban eltekintünk az általános inf jelölés használatától és feltesszük, hogy a keresett minimum létezik.

Gyakorlati alkalmazásokban a t_1, t_2, \dots, t_M időpontokban mintavételezett diszkrét $\mathbf{f} \in \mathbb{R}^M$ jelekkel dolgozunk, tehát $\mathcal{H} = \mathbb{R}^M$ a szokásos euklideszi skaláris szorzattal, a $\Theta_j(t, \mathbf{a})$ függvények megfelelő értékeit pedig a $\Theta(\mathbf{a}) \in \mathbb{R}^{M \times N}$ mátrixban tároljuk. Így a (2.2) probléma diszkrét megfelelőjét kapjuk [GolPer73]:

$$(2.3) \quad \min_{\mathbf{a}} \|\mathbf{f} - \mathbf{P}_{\Theta(\mathbf{a})}^N \mathbf{f}\| = \min_{\mathbf{a}} \|\mathbf{f} - \Theta(\mathbf{a})\Theta(\mathbf{a})^+ \mathbf{f}\|,$$

ahol $\Theta(\mathbf{a})^+$ jelöli a $\Theta(\mathbf{a})$ mátrix Moore–Penrose-féle pseudoinverzét. A minimalizálandó célfüggvényt *Variable Projection (VP)* funkcionálnak nevezzük, melynek számos gyakorlati alkalmazása ismert a telekommunikációban, robotikában, gépi tanulásban, adatelemzésben, differenciálegyenletek megoldásában, stb. [GolPer03]. Az elmúlt években a (2.3) módszert kiterjesztettem B-splineokra, racionális, és Hermite függvényekre, illetve többféle általánosítást is bevezettem. Ezeket mutatom be a következő pontokban.

2.1.1. Spline-ok illesztése az alappontok variálásával

Széleskörű alkalmazási lehetőségeinek (pl. számítógépes grafikában) köszönhetően a spline-ok szabad csomópontjainak optimális megválasztása a mai napig aktívan kutatott területnek számít. A state-of-the-art módszerek különböző megközelítéseket használnak az alappontok optimalizációjára, például genetikus algoritmusokat [YosHar03; UyaÜlk17], Lasso-típusú convex optimalizációt [YuaChe13], illetve kombinatorikus optimalizációt [KarGab97]. A [KovFek19] publikációmban egy új

adaptív projekciós megközelítést dolgoztam ki, melyben a spline-okat B-spline bázisban definiáltam:

$$B_{j,0}(t, \mathbf{a}) := \begin{cases} 1 & \text{ha } a_j \leq t < a_{j+1}, \\ 0 & \text{különben,} \end{cases}$$

$$B_{j,\ell}(t, \mathbf{a}) := \beta_{j,\ell} B_{j,\ell-1}(t, \mathbf{a}) + \gamma_{j,\ell} B_{j+1,\ell-1}(t, \mathbf{a}),$$

ahol

$$\beta_{j,\ell} = \frac{t - a_j}{a_{j+\ell} - a_j}, \quad \gamma_{j,\ell} = \frac{a_{j+\ell+1} - t}{a_{j+\ell+1} - a_{j+1}}.$$

Az adaptív projekciókhoz előre rögzítjük a spline-ok foksámát ℓ -et, a reprezentáció elemi függvényeit pedig a fenti B-spline-oknak választjuk, azaz $\Theta_j(t, \mathbf{a}) = B_{j,\ell}(t, \mathbf{a})$.

A csomóponti $\mathbf{a} = (a_0, a_1, \dots, a_{n-1})^T$ vektor (2.3)-beli optimalizálása nehéz feladat, ami az optimalizáció során egymáshoz közel kerülő alappontoknak köszönhető [Jup78]. Ezért itt az optimalizáció inicializálása különösen fontos. A [KovFek19] publikációmban nulladfokú spline approximációkat használva, az eltérést $\|\cdot\|_p$, ($p = 1, 2, \infty$) normákban minimalizálva dolgoztam ki különböző stratégiákat a kezdő alappontok megválasztására. Ezek a stratégiák helyesek abban az értelemben, hogyha a közelítendő függvény egy spline, akkor az inicializáció visszaadja az eredeti alappontokat. A valós jeleken és tesztfüggvényeken végzett szimulációk megmutatták, hogy a bemutatott módszerrel kapott kezdőpontokból indulva már 4 VP iterációt követően sokkal jobb optimumba jutunk, mint például egyenletes vagy véletlen pontrendszer esetén. A tesztek során a state-of-the-art algoritmusokhoz képest 12×-es gyorsítást sikerült elérni közel azonos hibahatár mellett. A B-spline-ok kompakt tartójú függvények, így a (2.3) egyenletben a $\Theta(\mathbf{a})$ mátrix oszlopaiban többségében nullák szerepelnek. Ezt kihasználva, például rikta adatstruktúrák segítségével tovább lehet gyorsítani az algoritmus implementációját, ami az [KovFek19] publikáció mellékleteként nyílt hozzáférésű kódként elérhető.

A kidolgozott nulladfokú spline approximációs módszerek nagyon hatékonyan számolhatók, ezért az inicializációs stratégiákkal a spline reprezentációk komplexitása, a csomópontok n száma is könnyen becsülhető. Ehhez egyre növekvő n -ekre kell kiszámolni az inicializációban szereplő nulladfokú spline approximációkat. Az így kapott alappontszám-approximációs hiba grafikon "sarokpontja" (maximális görbületű pontja) alapján meg lehet határozni a jelhez optimális csomópontok n számát. Megjegyzem, hogy a módszer nagyon hasonló elven működik, mint a Tikhonov-regularizációs algoritmusok regularizációs paraméterének becslésénél alkalmazott L-görbe módszer [Han92].

2.1.2. Paraméterezett Hermite-függvények

Ismeretes, hogy a klasszikus Hermite-polinomok $\{h_k \mid k \in \mathbb{N}\}$ ortogonális és teljes rendszert alkotnak az $(L_2(\mathbb{R}), \langle \cdot, \cdot \rangle_w)$ Hilbert-térben az alábbi skaláris szorzattal [Sze67]:

$$(2.4) \quad \langle h_k, h_\ell \rangle_w = \int_{-\infty}^{\infty} h_k(t) h_\ell(t) w(t) dt = \delta_{k\ell},$$

ahol $w(t) = e^{-t^2}$ a súlyfüggvényt jelöli. Ezek segítségével definiálhatók az alábbi Hermite-függvények:

$$\Phi_k(t) = h_k(t) / \|h_k\|_2 \cdot \sqrt{w(t)} \quad (k \in \mathbb{N}).$$

A w súlyfüggvénynek köszönhetően a Φ_k függvények nagyon gyorsan eltűnnek a $\pm\infty$ -ben, ezért praktikus szempontból tekinthetők kompakt tartójú függvényeknek. Emiatt kiválóan használhatók kompakt tartójú jelalakok reprezentációjára (pl. tüskék, akciós potenciálok, EKG jelek P, QRS, T hullámai, stb.). Hermite-függvények esetén azonban nincsenek szabad paramétereink, melyeket optimalizálni lehetne. Ennek a feloldására két módszert dolgoztam ki.

A [DózKov15] publikációban a $\lambda(t - \tau)$ affin argumentum transzformáción keresztül új szabad paraméterekkel, a λ dilatációval és a τ translációval, egészítettem ki az Hermite-függvényeket. Az így kapott $\Phi_k(t, \mathbf{a}) := \sqrt{\lambda} \Phi_k(\lambda(t - \tau))$ függvények öröklik az Hermite-függvények előnyös matematikai tulajdonságait (ortogonalitás, teljesség, "kvázi kompaktság"), ahol $\mathbf{a} = (\tau, \lambda)^T$ az optimalizálandó paraméterek vektorát jelöli. A számítógépes szimulációk során kiderült az is, hogy az így definiált függvénycsalád általában jól alkalmazható EKG jelek tömörítésére. A [KovFri20] munkámban racionális és B-spline alapú EKG reprezentációkkal is összehasonlítottam a módszert. A tesztek rámutattak arra is, hogy bizonyos speciális jelalakok (pl. abnormális szívütések, neuronok tüzelése, stb.) közelítéséhez nem elegendő az eredeti Hermite-függvényrendszer dilatációja és eltolása.

Az ábrázolható jelek osztályát a [KovBöc19] publikációban tovább bővítettem. Ehhez az Hermite-függvények w súlyfüggvényét módosítottam:

$$\begin{aligned} v_1(t, a) &:= u_1(t, a) \cdot w(t), & u_1(t, a) &:= (t - a)^2, \\ v_2(t, \mathbf{a}) &:= u_2(t, \mathbf{a}) \cdot w(t), & u_2(t, \mathbf{a}) &:= 1 / ((t - a_1)^2 + a_2^2) \\ && (t \in \mathbb{R}, \mathbf{a} \in \mathbb{R}^2). \end{aligned}$$

Az új v súlyfüggvényt v_1 és v_2 lineáris kombinációjaként kapjuk, melynek segítségével definiálhatók az új paraméterezett függvényrendszer elemei:

$$\Psi_k(t, \mathbf{a}) = q_k(t) / \|q_k\|_2 \cdot \sqrt{v(t, \mathbf{a})} \quad (k \in \mathbb{N}, t \in \mathbb{R}),$$

ahol $\{q_k \mid k \in \mathbb{N}\}$ a $\langle \cdot, \cdot \rangle_v$ skaláris szorzat szerinti ortogonális polinomokat jelöli. Így már fel lehet írni a (2.3) egyenletben definiált optimalizációs feladatot, melyben most a projekció legkisebb négyzetes hibáját minimalizáló v súlyfüggvényt, illetve annak paramétereit keressük. Az eredeti feladat gradiensek kiszámítása itt már elbonyolódik, ezért a numerikus optimalizáció implementációjához több egyszerűsítést is bevezettem. Többek között különböző nemlineáris feltételekkel leszűkítettem az alkalmas \mathbf{a} paraméterek halmazát, illetve egyszerűsítettem a célfüggvényt. Az algoritmus böngészőből futtatható implementációja a dolgozatokhoz készült szimulációkkal együtt elérhető a *codeocean* platformon [KovBöc18].

2.1.3. Racionális függvények pólusoptimalizációja

Az előző fejezetekben bemutatott reprezentációk optimalizálásánál csak a pontosságot vettük figyelembe, a komplexitást nem. Utóbbi, az $\mathcal{S}(\mathbf{a})$ altér N dimenziójával előre megbecsültünk, majd rögzítettünk. A [KovFri20] publikációban formalizálom a reprezentációk komplexitását amit együtt optimalizálok a pontossággal. Ehhez a (2.3) feladatot a következő módon általánosítottam:

$$(2.5) \quad r(\mathbf{a}, d) := \left\| f - P_{\Theta(\mathbf{a}, d)}^{N(d)} f \right\| + \Lambda(d),$$

ahol $d > 0$ a reprezentációkat indexelő egész szám, $\Lambda(d)$ a komplexitást mérő monoton függvény, $N(d)$ a projekció $\mathcal{S}(\mathbf{a}, d) := \text{span} \{ \Theta_j(\cdot, \mathbf{a}, d) \in \mathcal{H} : 0 \leq j < N(d) \}$ képterének dimenziója. A korábbi (2.3) megközelítéshez képest lényeges különbség, hogy a $P_{\Theta(\mathbf{a}, d)}^{N(d)}$ projekció nem csak az \mathbf{a} paramétervektortól, hanem a d komplexitástól is függ, aminek a pontosságát ellensúlyozza a Λ büntetőfüggvény. Ennek köszönhetően az $\mathcal{S}(\mathbf{a}, d)$ altér dimenziója csak akkor növekedhet, ha azzal jelentősen csökken az altértől vett távolság.

A (2.5)-ben adott funkcionál egy nagyon általános formalizálása a célfüggvénynek, amit konkrét alkalmazások esetén még definiálni kell. A [KovFri20] munkámban racionális ortogonális rendszerekre, az ún. Malmquist–Takenaka (MT) függvényekre implementáltam a $P_{\Theta(\mathbf{a}, d)}^{N(d)}$ projekciót és a Λ büntetőfüggvényt, ami egydimenziós jelek approximációjára alkalmazható. Ezzel az átfogalmazással egy kevert egészértékű optimalizálási feladathoz jutunk, amit egy híres részecske-raj alapú globális optimalizálási módszer (PSO [KenEbe95]) módosításával oldottam meg. Ez egyrészt, az eredeti algoritmus kiterjesztését jelenti a kevert egészértékű feladatra, másrészt, olyan feltételek beépítését, melyek a MT-függvények szabad paramétereit, az inverz pólusokat, a komplex egység körön belül tartják. Ezt a hiperbolikus geometria Poincaré-féle körmodelljében érvényes szerkesztések segítségével oldottam meg. Egy másik munkámban figyelembe vettem azt is, hogy EKG jelek esetén az említett inverz pólusok helye diagnosztikai információt hordoz. Ezért további, orvosi szempontból értelmezhető optimalizálási feltételeket is kidolgoztam a pólusok lokalizálására [Kov17b].

Megjegyzem, hogy ez a téma tekinthető a doktori tanulmányaim alatt megkezdett racionális függvényapproximációs módszerekkel kapcsolatos kutatásaim folytatásának is. Ehhez igazodva, a kidolgozott optimalizációs módszerek hatékonyságát egy konkrét gyakorlati alkalmazáson, EKG jelek tömörítésén teszteltem, ahol a cél az EKG jelek minél pontosabb közelítése a lehető legkevesebb adat felhasználásával. A [KovFri20; BogFri19] munkákban megmutattuk, hogy a módszer számos state-of-the-art algoritmusnál hatékonyabban tömörít, sőt egy nemrég megjelent szintén racionális függvényeket alkalmazó tömörítő módszer [TanZha19] eredményeit is sikerült megjavítani. A bemutatott optimalizációs algoritmusokat és a dolgozatokhoz készült szimulációkat böngészőből is futtatható interaktív formában tettem közzé a *codeocean* platformon [Kov17a].

2.2. Ritka reprezentációk

Az adaptív projekciós módszereknél az $f \in \mathcal{H}$ jelet egy nemlineárisan paraméterezett véges dimenziós $\mathcal{S}(\mathbf{a}) \subset \mathcal{H}$ alterbe képeztük, melyhez ortogonális vetítéseket alkalmaztunk, az alteret pedig egy $\{\Theta_j(\cdot, \mathbf{a}) : 0 \leq j < N\}$ bázisával adtuk meg. A ritka reprezentációk előállításának más megközelítései is vannak. Például a kérdéses N dimenziós alteret egy bővebb $G := \{\Theta_a : a \in \Gamma\}$ generátorrendszerével is megadhatjuk, ahol Γ egy index halmaz. Ebben az esetben az $f \in \mathcal{H}$ jelet (1.1) értelemben legjobban közelítő $\tilde{f} \in \mathcal{S}$ elem felírása nem egyértelmű, azaz többféleképpen előállítható G elemeinek lineáris kombinációjaként. A feladat az, hogy ezekből a reprezentációkból kiválasszuk a lehető "legritkébbat".

A ritka reprezentációk keresése különböző típusú matematikai optimalizálási problémáknak feleltethető meg. Az egyik ilyen az alábbi kombinatorikus optimalizációs feladat:

$$(2.6) \quad \min_{|\Lambda|=N} \min_{\mathbf{c}} \left\| f - \sum_{a \in \Lambda} c_a \Theta_a \right\|_2,$$

ahol $\Lambda \subset \Gamma$. A pontos megoldás kiszámítása NP-nehéz [Tro04a], amit különböző heurisztikus módszerekkel [MalZha93; CheDon01] szokás becsülni.

Az (2.6) megközelítésben előre adott, hogy hány elemű lineáris kombinációval közelítjük f -et. Ha még ezt sem tesszük fel, akkor a következő problémához jutunk:

$$(2.7) \quad \min_{\mathbf{c}} \left\{ \left\| f - \sum_{a \in \Gamma} c_a \Theta_a \right\|_2 + \lambda \cdot \|\mathbf{c}\|_0 \right\},$$

ahol $\|\cdot\|_0$ az argumentum nem nulla koordinátáinak számát jelöli, a büntetőtag $\lambda > 0$ súlyát pedig regularizációs paraméternek nevezzük. Ez szintén NP-nehéz feladat [Tro04b], amit a büntetőtag convex $\|\cdot\|_1$ relaxációjával szokás kezelni.

A [Kov18] publikációmban ritka racionális reprezentációk előállítását vizsgáltam, melyhez mohó [MalZha93] és Lasso-típusú [Tib09] algoritmusokat alkalmaztam. Általában véve, a ritka reprezentációs elméleti eredmények (pl. hibabecslések) bizonyításában komoly szerepet játszik a függvényrendszer koherenciája, mely a reprezentáció felírásához használt elemi függvények közötti hasonlóságot méri. A [Kov18] munkámban meghatároztam a racionális generátorrendszerek koherenciáját, illetve a speciális, egy pólust tartalmazó Laguerre-rendszerek unióján keresztül empirikusan is megvizsgáltam a kapott formulák helyességét.

3. fejezet

Gépi tanulás

A 2010-es évek elejéig, a jelfeldolgozásban felmerülő problémák megoldását a matematikai modell alapú szemlélet dominálta. Ezek a modellek a jelfeldolgozási problémákat az 1. fejezetben bemutatott egyszerűbb lineáris, illetve nemlineáris approximációs feladatokra vezették vissza, melyek elméleti háttére jól ismert. Így interpretálható matematikai modelleket kapunk, melyekbe könnyen beépíthetők az adott probléma jellemzői, például simasági, statisztikai feltételek, a közelítő altér topológiája, vagy a zaj típusa. Bonyolultabb problémák esetén azonban ez a fajta szigorú matematikai modell alapú megközelítés követhetlenné válik. Ezekre a kihívásokra jelenthetnek megoldást a gépi tanulás során felépített adatvezérelt modellek.

A 2.1 pont modell alapú megközelítéseihez hasonlóan, a gépi tanuló algoritmusok is a jel egy hatékony, azaz pontos és egyszerű, reprezentációját állítják elő. A tanítás során nemlineárisan paraméterezett függvények egy tág $\Theta(\mathbf{a}) := \{\Theta_a : a \in \Gamma\}$ osztályát optimalizáljuk, melyek segítségével hatékonyan ábrázolhatók a tanításhoz használt jelek. Ha a tanítóhalmaz lefedi a vizsgált jelenség eloszlását, akkor az algoritmus a korábban nem látott méréseken is az elvárt helyes eredményt adja.

A gépi tanulás egy kitüntetett területe a mesterséges mély neurális hálózatok, melyekben rétegekbe rendezett mesterséges neuronok kommunikálnak egymással. Az előrecsatolt háló az egyik leggyakoribb ilyen struktúra, melyben a neuronok egymásutáni rétegekbe szerveződnek:

$$(3.1) \quad \tilde{f} = NN_{\mathbf{a}}f = \left(\Theta_{a^{(N)}}^{(N)} \circ \dots \circ \Theta_{a^{(j)}}^{(j)} \circ \dots \circ \Theta_{a^{(2)}}^{(2)} \circ \Theta_{a^{(1)}}^{(1)} \right) f,$$

ahol $f \in \mathcal{H}$ a bementi jel, \tilde{f} a háló kimenete, $\Theta_{a^{(j)}}^{(j)}$ a j -ik rétegben végrehajtott leképezést, $a^{(j)}$ a réteg paramétereit, az \mathbf{a} szimbólum pedig az összes optimalizálendő $a^{(j)}$ paraméter halmazát jelöli. A rétegeket leíró $\Theta_{a^{(j)}}^{(j)}$ műveletek lehetnek lineáris leképezések (pl. konvolúciók), nemlineáris aktivációs függvények (pl. szigmoid, hiperbolikus tangens), stb. A háló kimenete \tilde{f} az eredeti f jel egy nemlineáris reprezentációjaként értelmezhető, ám a 2. fejezet additív megközelítéseivel ellentétben, (3.1) a jel egy kompozív ábrázolása. A kompozív approximációs eljárás is

helyes abban az értelemben, hogy minden a $K \subset \mathbb{R}^p$ kompakt halmazon értelmezett folytonos $g : K \rightarrow \mathbb{R}^q$ függvény tetszőleges pontossággal megközelíthető egy olyan $NN_{\mathbf{a}}$ típusú leképezéssel, melynek belső (rejtett) $\Theta_{a^{(j)}}^{(j)}$ ($j = 2, \dots, N-1$) rétegeiben az aktivációs függvények nem polinomok, vagyis a nem polinomiális aktivációs függvényű $NN_{\mathbf{a}}$ leképezések \mathcal{N} osztálya sűrű a $(C(K), \|\cdot\|_{\infty})$ Banach térben [Cyb89; Hor91; KidLyo20].

A gépi tanulás tekinthető egy többdimenziós interpolációs/regressziós feladatnak is, melyben a tanító halmaz bemeneti-célérték relációi az interpolációs/regressziós feladat alappont-érték párjainak felelnek meg. A gyakorlatban előforduló interpolációs és regressziós feladatokban a közelítendő $g : K \rightarrow \mathbb{R}^q$ függvényt nem ismerjük, csak mérni tudjuk, a méréseket pedig (felügyelt tanítás esetén) pont-érték párokkal adjuk meg:

$$(3.2) \quad (f_i, g_i) \quad (i = 1, 2, \dots, M),$$

ahol $f_i \in K \subset \mathbb{R}^p$ és $g_i \in \mathbb{R}^q$. A mért érték, azaz a háló elvárt kimenete egy adott f_i mérési pontban, alkalmazástól függően lehet egy $\mathbf{g}_i \in \mathbb{R}^q$ vektor pl. regresszió esetén, vagy az osztályok $g_i \in \mathbb{N}$ címkéje osztályozási feladatokban, de ábrázolhat $g_i \in [0,1]^q$ valószínűségeket is. A tanítás során az adott problémához illeszkedő célfüggvényt minimalizáljuk az \mathbf{a} paraméterek szerint a (3.2) tanítóadatokon. Regressziós feladatokban például gyakran az átlagos négyzetes eltérés (MSE), míg osztályozási feladatokban a bináris keresztentrópia (BCE) minimumát keressük:

$$(3.3) \quad J_{MSE}(\mathbf{a}) := \frac{1}{M} \sum_{i=1}^M \|g_i - \tilde{g}_i\|_2^2,$$

$$(3.4) \quad J_{BCE}(\mathbf{a}) := -\frac{1}{M} \sum_{i=1}^M (g_i \log \tilde{g}_i + (1 - g_i) \log(1 - \tilde{g}_i)),$$

ahol $\tilde{g}_i = NN_{\mathbf{a}}f_i$ a háló f_i bemenetre adott válasza.

A neurális hálók struktúrája rendkívül moduláris, melynek számos elemét (pl. célfüggvény, aktivációs függvények, rétegek, stb.) a felhasználás helyétől függően lehet definiálni. Ez egy sok szempontból általánosabb keretrendszer, mint az 2. fejezetben bemutatott approximációs eljárások. Sőt, a J_{MSE} célfüggvény minimalizálásával, $g_i = f_i$ és $M = 1$ speciális (egy mintás autoenkóder) esetben tulajdonképpen a 2.1 pont approximációs feladatait oldjuk meg, csak itt projekciók helyett nemlineáris függvények kompozíciójával állítjuk elő az f jelet közelítő \tilde{f} elemet. Megjegyzem, hogy az említett tartalmazástól függetlenül mindkét szemléletnek van létjogosultsága. A mély hálók kimenete ugyanis egy többszörösen összetett nemlineáris függvény helyettesítési értékeként fogható fel, amiben rengeteg szabad paramétert kell optimalizálni. Ez egyrészt nagyon számításigényes feladat, másrészt a háló bemenetére adott válasz nehezen értelmezhető. A 2.1 pont projekciós megközelítéseivel szemben az ilyen mély hálós modelleknek nincs konkrét fizikai jelentésük, tulajdonképpen fekete doboznak tekinthetők. Ez viszont nem elfogadható az olyan alkalmazásokban, melyekben indokolható döntéseket kell hozni (pl. biológiai jelfeldolgozásban, autonóm járműirányításban). A következő két fejezetben,

az említett problémák megoldásához olyan tanuló algoritmusokat fejleszttek, melyekben a matematikai modell alapú szemlélet és az adatvezérelt modellek előnyeit kombinálom.

3.1. Hibrid tanuló algoritmusok

A matematikai alapú és az adatvezérelt modellek ötvözésének két fő motivációja a teljesítmény növelése és az algoritmusok futásidejének csökkentése. Ezek kritikus tényezők számos alkalmazásban és gyakran csak egymás rovására javíthatók. A tomográfiás képalkotás is egy ilyen felhasználási terület, melyben a cél a vizsgált objektum 3D-s rekonstrukciója. Ehhez a tárgyat valamilyen sugárzással átvilágítják, amelyet az anyag nagyrészből átenged, de kisrészben el is nyel és így az elnyelődés mértékéből következtetni lehet az objektum belső szerkezetére. A tomográfiás képalkotási eljárásokban a térbeli tárgyak struktúrája azok síkmetszeteiből rekonstruálható, melyhez több különböző irányból felvett egydimenziós mérést kell végezni. Az analitikus rekonstrukciós módszerekben a \mathbf{b} mérések és az anyag lokális \mathbf{x} elnyelődési együtthatói között egy $\mathbf{Ax} = \mathbf{b}$ típusú lineáris kapcsolatot feltételezünk. A fizikai folyamatot modellező \mathbf{A} lineáris transzformáció általában rosszul kondicionált, ezért az \mathbf{x} vektort csak valamilyen regularizációs módszer segítségével állíthatjuk elő:

$$(3.5) \quad \tilde{\mathbf{x}} = \arg \min_{\mathbf{x}} \{ \|\mathbf{b} - \mathbf{Ax}\|_2 + \lambda \cdot \Omega(\mathbf{x}) \},$$

ahol Ω egy alkalmasan választott büntetőfüggvény (pl. $\|\mathbf{x}\|_p$, $p = 1,2$), amit a $\lambda \geq 0$ regularizációs paraméter súlyoz. A fenti problémát síkmetszetenként különböző \mathbf{b} jobb oldalakra és néha különböző \mathbf{A} mátrixokra kell megoldani, amit az iteratív rekonstrukciós eljárások határértékben állítanak elő. A felbontás növelésével nemcsak a síkmetszetek száma növekszik, de a (3.5) egyenletben szereplő \mathbf{x} , \mathbf{b} vektorok és az \mathbf{A} mátrix dimenziója is. Ezért magas felbontású képek esetén az iteratív rekonstrukció nagyon számításigényes eljárás. A rekonstrukció minőségét erősen befolyásolja a mérést modellező \mathbf{A} mátrix konstrukciója, illetve a mérések és az elnyelődések között feltételezett lineáris kapcsolat is, ami csak közelíti a valóságban lezajló mérési folyamatot. Ezekre a kihívásokra jelentenek megoldást a gépi tanuló algoritmusok, melyek a kérdéses fizikai jelenségek nemlineáris természetét is képesek modellezni, ezáltal javítva a rekonstruált képet, csökkentve a zaj és a képalkotásból származó műterméket. Az adatvezérelt megközelítés másik előnye, hogy a betanított neurális hálók kiértékelése nagyon gyors, ami valós idejű rekonstrukciót tesz lehetővé [WilNoë19], köszönhetően az erősen párhuzamosítható GPU implementációknak.

A sok paramétert tartalmazó mély neurális hálók tanítása nehéz feladat, amihez rengeteg adatra van szükség. Számos tomográfiás alkalmazás esetén (CT, röntgenfelvétel, anyagvizsgálat, stb.) a mérési adatok begyűjtésének magas költségei vannak, például szükség van mesterségesen legyártott fizikai fantomokra, önkéntesekre, orvosi szakértőkre, ami megnehezíti a gépi tanuló algoritmusok alkalmazását.

Ebben a fejezetben olyan hibrid mély hálós rekonstrukciós eljárásokat dolgozok ki, melyekben a hálók méretét és a tanítás adatéhségét analitikus modellekkel generált szintetikus adatokkal és az (3.5) egyenlethez hasonló iteratív algoritmusokkal csökkentem.

3.1.1. Termografikus képrekonstrukció

A tomográfiás vizsgáltokban az átvilágításhoz különböző típusú sugárzás használható, ettől függően más és más képalkotó modalitások léteznek. Számítógépes tomográfia (CT) esetén például röntgensugarak elnyelődéséből lehet következtetni az anyag belső szerkezetére, de léteznek ultrahangon, illetve hősugárzáson alapuló modalitások is. Utóbbit termografikus tomográfiának nevezzük, melynek számos gyógyászati és ipari alkalmazása van. Az egyik ilyen a roncsolásmentes anyagvizsgálat, melyben a cél a gyártási folyamat vagy a használat során keletkezett szerkezeti hibák (légbuborék, törés, repedés, stb.) felismerése. Ehhez az vizsgált objektumot kívülről (pl. hőlámpával, indukcióval) felhevítik a felszíni hőmérsékletváltozást pedig infravörös kamerával rögzítik. A fizikai jelenséget a hővezetés differenciálegyenlete írja le:

$$(3.6) \quad \left(\nabla^2 - \frac{1}{\alpha} \frac{\partial}{\partial t} \right) T(\mathbf{r}, t) = -\frac{1}{\alpha} T_0(\mathbf{r}) \delta(t),$$

ahol α az anyag hővezetési tényezője, δ a Dirac-delta függvény, $T(\mathbf{r}, t)$ jelöli az anyag hőmérséklet eloszlását az \mathbf{r} térben és t időben, $T_0(\mathbf{r})$ pedig a keresett kezdeti hőmérséklet eloszlás. A fenti differenciálegyenlet megoldására számos numerikus módszer ismert, melyekkel ebben a dolgozatban nem foglalkozunk.

Burgholzer és szerzőtársai 2017-ben egy új típusú megközelítést publikáltak [BurTho17], melyben (3.6) direkt megoldása helyett egy kétlépéses rekonstrukciót dolgoztak ki. Az első lépésben a termografikus képrekonstrukciós problémát egy lineáris transzformáción keresztül fotoakusztikus képrekonstrukciós problémává konvertálták, amit a második lépésben immár a hullámterjedést leíró differenciálegyenletek megoldásával lehet kezelni. A módszer előnye, hogy a második lépésben alkalmazott módszereken keresztül figyelembe lehet venni a keresztirányú hőterjedést és így magasabb jel-zaj arányú (SNR) rekonstrukciót tudunk elérni. 2018-ban ezekbe a kutatásokba kapcsolódhattam be én is a linzi Johannes Kepler Egyetemen.

A kétlépéses rekonstrukció első fázisában egy (3.5) típusú regularizációt kell végrehajtani, melyben az \mathbf{A} mátrix explicit képlettel számolható, mérete pedig elhanyagolhatóan kicsi (általában pár ezerszer pár ezres nagyságrendű) az összes felszíni méréshez képest. A második lépésben is egy (3.5) típusú optimalizációs problémát kell megoldani, az ehhez tartozó \mathbf{A} mátrixot viszont csak közelítőleg ismerjük, aminek a mérete is sokkal nagyobb (a rekonstruálandó kép vagy térfogat felbontásával arányos). Ezért a [KovLeh20a] munkában a második regularizációs lépést egy U-Net architektúrájú neurális hálóval helyettesítettem. Az U-Net egy speciális autoenkóder struktúra, melyben távoli (enkóder-dekóder) rétegeket

összekötő kapcsolatok is vannak, azaz például (3.1)-ben az első $\Theta_{a(1)}^{(1)}$ réteg kimenete nemcsak a második, hanem további $\Theta_{a(j)}^{(j)}$, $j > 2$ rétegek bemenete is lehet. Az U-Net hálókat eredetileg orvosi és biológiai képszegmentáláshoz fejlesztették, melyekben kevés tanítóadat áll rendelkezésre [RonFis15].

A [KovLeh20a] publikációban megmutattuk, hogy az U-Net előnyös tulajdonságai kihasználhatók a roncsolásmentes anyagvizsgálatban is. Ez tulajdonképpen egy bináris osztályozási feladat, melyben a rekonstruált termografikus képeken a defekteket egyesek, a normális részeket pedig nulla értékű pixelek jelzik. Csakúgy, mint az orvosi és biológiai képszegmentálási feladatokban, a roncsolásmentes anyagvizsgálatban is költséges valódi annotált mérési adatokat rögzíteni, amihez mesterséges fizikai fantomokra van szükség. Az U-Net alkalmazásával persze csökkenthető a tanító adat mennyisége, de ez még nem elegendő. Ezért a tanító és a tesztalmaz elemeit az (3.6) differenciálegyenlet adiabatikus megoldásaival generáltuk. Ez azt jelenti, hogy a szimulációk során a testfelületen átmenő hőáramot nullának tekintettük. Az adiabatikus peremfeltétel alkalmazása leegyszerűsíti a hővezetés egyenlet megoldását, melynek segítségével egy tetszőlegesen generált T_0 hőmérséklet eloszláshoz szimulálni tudjuk a felszínen mért hőmérséklet eloszlást $T(0, t)$. Ennek az egyszerű modellnek a használatával egy 40 000 hőképet tartalmazó adatbázist generáltunk [KovLeh20c], melyen két U-Net variánst tanítottunk. A kétlépéses [BurTho17] sémát követve, a hőképeket először egy (3.5)-beli regularizációs módszerrel akusztikus jelekké konvertáltuk, amit továbbítottunk a neurális háló bemenetére. Az első modell alapú regularizációs lépés egy dimenziócsökkentésnek felel meg, melyben további matematikai feltételek (nemnegativitás és ritkaság) is kihasználhatók.

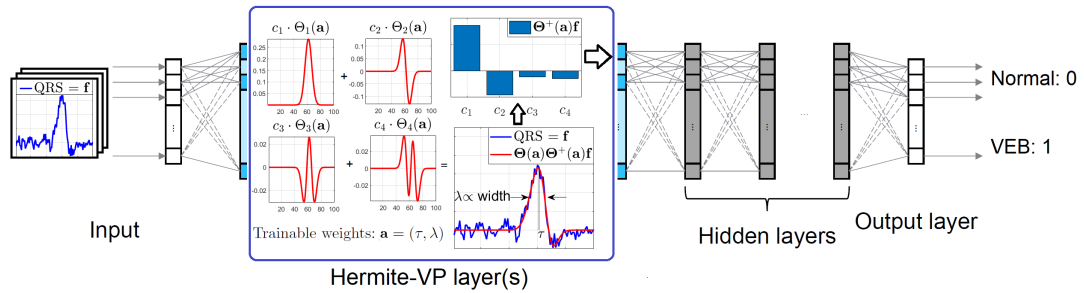
A [KovLeh20b] cikkben 100 000-esre bővítettük a szintetikus adatbázist, majd a hibrid tanítást összehasonlítottuk az egylépéses *end-to-end* megközelítéssel is, ahol a hálókat közvetlenül a felszíni méréseken tanítottuk (itt tehát nincs közbevetett matematikai transzformáció). A kutatás során kiderült, hogy az egylépéses rekonstrukció pontossága egy adott ponton túl (30dB SNR) nem javítható, azaz a tisztább felszíni mérések nem eredményeznek jobb rekonstrukciót. A hibrid módszer esetén viszont a felszíni mérések javításával a rekonstrukció hibája is szigorúan monoton csökkent. Megmutattuk, hogy a szintetikus adatokon betanított U-Net variánsok jól alkalmazhatók valós méréseken is. A tesztek során a rekonstrukciót különböző szögállású méréseken és eltérő zajszintek mellett is megvizsgáltuk, melyek során kiderült, hogy a kidolgozott hibrid tanítási módszerrel nemcsak pontosabb, de robusztusabb algoritmust kapunk, mint a state-of-the-art eljárások.

Az áttekintett munkákban a rekonstruált kép hibáját átlagos legkisebb négyzetes eltérésben (MSE) mértük, amihez szükség volt ismerni az eredeti képet. Ez a gyakorlatban nem áll rendelkezésre, csak a rekonstrukció során kapunk róla becslést. A [LehGal21] cikkben két könnyen számolható hibabecslést dolgoztunk ki, melyek a referencia kép ismerete nélkül becsülik a rekonstruált kép hibáját. Így becsülhető a rekonstruált képen látszó defektek bizonytalansága, de a metrikák jól alkalmazhatók annak eldöntésére is, egy konkrét feladatban, hogy melyik betanított háló struktúra eredményét fogadjuk el helyesnek.

3.2. Modell-vezérelt neurális hálók

A matematikai modell alapú szemlélet és az adatvezérelt modellek előnyeinek megtartását célozzák a modell-vezérelt neurális hálók. Ez az ún. Explainable AI (XAI) kutatások egy kitüntetett témaköre, melyben a neurális háló struktúráját valamilyen matematikai probléma paraméterezett megoldásaként keressük. Példaként említhető a ResNet [HeZha16], illetve az ODENet [CheRub18], melyekben az egyes rétegek kimenete paraméterezett közönséges differenciálegyenletek megoldásaként értelmezhető [ChaMen18]. A gyakorlatban felmerülő optimalizációs problémák megoldásához általában szükség van alkalmazásspecifikus ismeretekre, melyek matematikai formalizálása, a priori információk (simaság, konvexitás, ritkaság, stb.) megfogalmazása nem könnyű feladat. Ez motiválta az OptNet [AmoKol17] hálóstruktúrák létrehozását, melyekben a rétegek paraméterezett kvadratikusan programozási (QP) feladatokat reprezentálnak. Az egyes rétegek paraméterei a QP feladat egyenlőség és egyenlőtlenség feltételei, melyeket a modell alapú megközelítésekkel ellentétben nem kell explicit definiálni, a háló éppen ezeket tanulja meg. QP helyett a (3.5) típusú optimalizációs feladat is beépíthető a neurális hálók rétegeibe az ún. DeepUnfolding [HerRou14] eljáráson keresztül, mely a (3.5)-öt megoldó iterációs módszerek lépéseit rétegekbe fejtí. Az így kapott neurális hálók rétegaparaméterei a kifejtett numerikus módszer paramétereivel egyeznek meg (pl. a λ regularizációs paraméterrel). Bizonyos jelfeldolgozási problémák inspirálták a Wiener-[Law13], és Hammerstein-típusú [YuPen14] neurális hálókat, melyek dinamikus lineáris blokkok és statikus nemlineáris blokkok kompozíciói. A dinamikus blokk általában egy lineáris időinvariáns (LTI) rendszert, a statikus pedig egy betanított neurális hálót reprezentál.

A [KovBog21b] publikációmban a fenti modell-vezérelt neurális hálótól eltérő, teljesen új konstrukciót, ún. adaptív projekciós hálókat (VPNet) dolgoztam ki. A szükséges elméleti háttér a 2. fejezetben bemutatott VP funkcionálok adják, melyek segítségével szeparálható nemlineáris legkisebb négyzetes problémákat lehet megoldani, ahol a paraméterek lineáris és nemlineáris csoportokra bonthatók. A módszert olyan problémákban célszerű alkalmazni, melyekben kevés nemlineáris változótól függő sok lineáris paramétert kell meghatározni. Számos jelfeldolgozási probléma hasonló jellegű, például a jelek tömörítése, reprezentációja, illetve a jelekből történő jellemzők (feature-ök) kinyerése. Az adaptív projekciós hálók, a konstrukciónak köszönhetően, örökölik a VP operátorok jó tulajdonságait, ami értelmezhető paramétereket, illetve kompakt, áttekinthető háló struktúrákat eredményez. A VPNet alkalmazható regressziós és osztályozási feladatok megoldására is. Utóbbit, szintetikus és valós adatokon is teszteltük, melyhez a 2.1.2 alpontban bemutatott paraméterezett Hermite-függvényekkel definiált VPNet-et használtuk. Az így kapott hálóstruktúrákat a 3.1 ábra szemlélteti. Látható, hogy az első VP réteg tanítható paraméterei az Hermite-függvények τ eltolása és λ dilatációja. Ezt a hálót a [KovBog21b] cikkben EKG jelek QRS komplexumának bináris osztályozására használtuk, tehát annak eldöntésére, hogy az adott QRS normális vagy abnormális szív működést jelez-e. A 3.1 ábrán jól látszik, hogy az első VP réteg tulajdonképpen



3.1. ábra. VPNet architektúra egy realizációja Hermite-függvényekkel.

a QRS komplexumok szélességét és időbeli pozícióját tanulja meg és a (2.1) egyenlet c_j együtthatóit továbbítja a következő rétegnek. A VPNet első rétege tehát egy automatikus jellemzőkinyerést végez, ám más hasonló (pl. konvolúciós) hálóstruktúrákkal ellentétben, a jellemzők és a tanított paraméterek orvosi szempontból is értelmezhetők maradnak. Továbbá a VP réteg a jel egy ritka reprezentációját állítja elő, melyet úgy értünk el, hogy az osztályozáshoz használt (3.4)-beli BCE célfüggvényt a reprezentáció hibáját büntető taggal egészítettük ki. Így a teljesen kapcsolt, illetve a konvolúciós neurális hálókhoz képest sokkal kevesebb paraméterrel értük el ugyanazt az osztályozási pontosságot. A [KovBog21b] cikkel együtt a háló Python implementációját is publikáltuk, mely alapján az összes teszt és szimuláció reprodukálható [KovBog21a].

A VPNet hálók kutatásán jelenleg is aktívan dolgozom, melyben számos nyitott kérdés vár megválaszolásra: hogyan lehet a konstrukciót általánosítani más gépi tanuló algoritmusokra (pl. SVM, K-means); milyen problémák megoldására érdemes alkalmazni a VPNet-et; ezekhez milyen függvényrendszer illeszkedik; mennyire jól/rosszul kondicionált a VP réteg egy adott függvényrendszer esetén? Az említett kutatási feladatokat az idén elnyert MTA Bolyai János Kutatási Ösztöndíj keretében végzem.

4. fejezet

Alkalmazások

Az előző fejezetekben kidolgozott modelleket, algoritmusokat konkrét gyakorlati problémák megoldására is alkalmaztam. Ez gyakran interdiszciplinális jellegű komplex kutatásokhoz vezetett, amihez nem csak informatikai, illetve matematikai tudásra, de alkalmazásspecifikus ismeretekre és különböző területek szakembereivel való együttműködésre is szükség volt. A következő pontokban először összefoglalom az EKG és EEG jelfeldolgozásban elért eredményeimet, melyekben a linzi Kepler Egyetemi Kórház kardiológusaival működhettem együtt. Megjegyzem, hogy az EKG és EEG jelekhez kapcsolódó kutatásaimat még a doktori tanulmányaim alatt az ELTE IK Numerikus Analízis Tanszékén kezdtem, a habilitációs dolgozatban viszont csak a védést követően elért eredményekre koncentrálok. A második alfejezetben a rendszeridentifikációhoz kapcsolódó kutatásaimat tekintem át, melyekben a linzi Johannes Kepler Egyetem (JKU) Jelfeldolgozó Intézetének mérnökeivel és a Felső-Ausztriai Alkalmazott Tudományok Egyetemének (FHOÖ) Josef Ressel Intézetének fizikusaival dolgozhattam együtt.

4.1. EKG, EEG jelfeldolgozás

A szívizom által generált elektromos feszültség információt hordoz a szív működéséről, mely a szív- és érrendszeri megbetegedések diagnózisában kulcsfontosságú szerepet játszik. A szív elektromos aktivitását a testre helyezett elektródákon keresztül lehet mérni. Két elektróda között az idő függvényében mért elektromos potenciálkülönbséget nevezzük EKG jelnek. A szív ciklikus működésének köszönhetően ezek kváziperiodikus jelek, vagyis kisebb változásoktól eltekintve ugyanazok a jelalakok ismétlődnek egymás után. Egy perióduson belül általában három jól kivehető csúcsot kell látnunk, melyek a szinusz csomóból induló akciós potenciál terjedését reprezentálják a szív egyes részeiben. A P hullám a pitvari izomzat kisülését, a QRS hullám a kamraizomzat összehúzódását, a T hullám pedig ennek az elektromos újratöltődését jelzi. Ezen kívül más, diagnosztikai szempontból fontos részei is vannak az EKG jelnek (pl. ST-segmens, U-hullám, stb.), de ezek diszkussziójától ebben a dolgozatban eltekintünk.

Attól függően, hogy az emberi test mely pontjaira helyezzük az elektródákat,

a P, QRS, T hullámok karakterisztikája más és más lehet, de egy elvezetésen belül normális esetben nincs számottevő alaki változás. Elvezetésenként a három hullám modellezéséhez rögzíthetünk egy-egy adaptív függvényrendszert a 2. fejezetből, amit a szabad paraméterek optimalizációjával hozzáigazíthatunk az apróbb morfológiai változásokhoz. A paramétereken keresztül követhetjük az EKG jelek alaki változásait, amiből következtetni lehet a szív állapotára, abnormális szív működésre, ritmuszavarokra, stb. Az ELTE Numerikus Analízis Tanszékén régóta folynak EKG jelfeldolgozással kapcsolatos kutatások, különös tekintettel azok racionális függvényekkel való modellezésére [FriLóc12]. Ezekben a kutatásokban a szívizom összehúzódnás által gerjesztett elektromos erőter eredőjét két elsőrendű és egy másodrendű pólussal definiált elemi racionális függvények lineáris kombinációjaként reprezentálák. A kardiológiában használt fogalmak (a szív elektromos főtengelye) és szabályok (Einthoven-féle törvény) könnyen magyarázhatók a racionális modell segítségével. Sőt, ez a síkbeli modell helyesen adta vissza a végtagi elvezetéseken mért jelek QRS komplexumát, melyek az erőter különböző irányok szerinti vetületeiként foghatók fel. A QRS komplexum előállítása jól tükrözte a valódi EKG jelek különböző elvezetésinek karakterisztikáját is, a jel többi részét, a P és T hullámokat, viszont nem. A 2D modell a racionális kvaternió-változós függvényeken keresztül [PapSch18] kiterjeszthető 3D-be, melyben a pólusokat elvezetésektől független, a páciensre jellemző szabad paraméterként lehet kezelni. Így már nemcsak a QRS, de a P és T hullámok is helyesen ábrázolhatók az egyes elvezetéseken. Valós EKG jelek esetén a modell illesztése, a paraméterek optimalizációja jelenleg is aktív kutatási területnek számít.

A racionális EKG modellekben a jel egészét közelítettük, mely jól alkalmazható adattömörítésre, zajszűrésre, az orvosi szempontból fontos P, QRS, T hullámok automatikus szétválasztására azonban nem. Ezért a 2.1.2 alpontban bemutatott paraméterezett Hermite-függvények segítségével új EKG modelleket definiáltam [KovBöc17], melyekben az említett hullámokat külön-külön közelítettem. A modell hullámonként két szabad paramétert, egy τ eltolást és egy λ dilatációt, tartalmaz, összesen tehát hat paramétert kell optimalizálni szívütésenként. Az optimalizáció végén, az eltolás megadja a hullámok helyét, a dilatáció pedig a szélességét, amiből meg lehet határozni a P, QRS, T hullámok kezdő és végpontjait, illetve az ezekhez tartozó diagnosztikai jellemzőket (QT/PR intervallum és PR/ST szakaszok hossza, stb.). A [KovBöc17] publikációban a Hermite-féle modellt EKG jelek szegmentálására alkalmaztam, a paraméterek optimalizációjához pedig orvosi szempontból értelmezhető lineáris feltételeket vezettem be. Az algoritmust a Physionet QT adatbázisán [GolAma00] teszteltük, melyet state-of-the-art módszerekkel is összehasonlítottunk. Az tesztek eredményei nagyon biztatóak voltak, ám észrevettük, hogy erős alapvonalvándorlású és alapvonalugrású EKG görbék esetén a modell nem tökéletes. A [BöcKov21] publikációban a problémás jelalakokat spline-okra és szigmoid függvényekre ültetett Hermite-függvényekkel modelleztem. Ezzel a kiterjesztéssel ugyan elveszítjük az Hermite-féle rendszer ortogonalitását, viszont az EKG jelek egy sokkal szélesebb osztályát tudjuk reprezentálni. Sőt a szegmentáció mellett, a spline-oknak köszönhetően el is tudjuk távolítani a jelből

a diagnoszikai szempontból zajnak minősülő alapvonalat. A paraméterek optimalizációjához további, orvosi szempontból releváns (pl. hullámok egymáshoz való viszonya, pozitivitás), nemlineáris feltételeket dolgoztam ki. Megjegyzem, hogy ez egy tisztán modell alapú szegmentáló algoritmus, melyhez nem használtunk neurális hálókat. Azonban a kidolgozott EKG modell csak a 2.1.2 alpontban bemutatott adaptív projekciókon alapszik, ezért a 3.2 pont mintáját követve könnyen hozzákapcsolható és együtt tanítható egy neurális hálóval. Az így kapott speciális VPNet első rétege a szegmentálandó QRS, T és P hullámok helyét és szélességét tanulná meg. Egy ilyen konstrukció tényleges implementációja azonban számos kérdést felvet, ami további kutatásokat igényel.

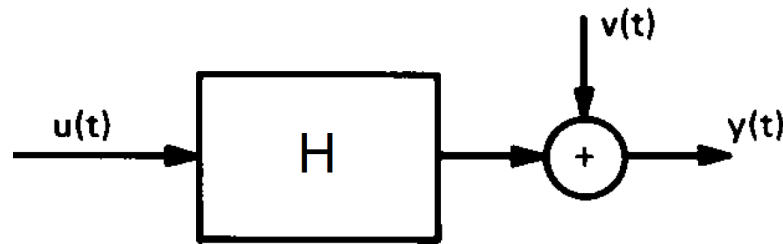
A szegmentálás az orvosi diagnózis felállításának első, előfeldolgozó lépése, melyet követhet például a szívütések osztályozása. A [DózBog19] munkában a 2. fejezetben bemutatott adaptív projekciók segítségével kinyert jellemzők felhasználásával a szegmentált EKG jel szívütéseit 5 (1 normális + 4 abnormális) osztályba soroltuk. Az adaptív projekciók szabad paramétereit páciensenként, az osztályozástól függetlenül optimalizáltuk, majd a rögzített nemlineáris paraméterekkel kiértékelt projekciók együtthatóit egy tartóvektorgépnek (SVM) továbbítottuk. A módszereket a PhysioNet MIT-BIH Arrhythmia Adatbázison [GolAma00] teszteltük, mely több mint 100 000 szívütést tartalmaz. A cikkben megvizsgáltuk a spline-, racionális- és Hermite-reprezentációkon külön-külön tanított SVM-ek teljesítményét, és az azok kombinálásával kapott ún. ensemble módszerek hatékonyságát is. Az így kapott algoritmusok átlagosan 90% feletti osztályozási pontosságot értek el, amivel számos state-of-the-art módszert sikerült magunk mögé utasítani.

A biológiai jelfeldolgozás egy másik fontos területe az elektroencefalográfia (EEG), mely az agyi idegsejtek összegzett elektromos aktivitását reprezentálja. Az EEG-t a fejre helyezett elektródákon keresztül lehet mérni, melyet számos tényező (haj, fejbőr, szemmozgás, stb.) befolyásol, ami megnehezíti az alkalmazáspecifikus algoritmusok kidolgozását. Ennek ellenére, az EEG jelfeldolgozásnak rengeteg gyakorlati alkalmazása van például a betegségek felismerésében, alvásfázis detektálásban, protézisek vezérlésében, melyek közül én az epilepsziás rohamok detektálásával foglalkoztam [SamKov17]. Ez a doktori tanulmányaim alatt megkezdett kutatások folytatása, melyben az EEG jelek 1 másodperces szegmenseihez tartozó racionális komponensek felhasználásával definiáltunk tulajdonságvektorokat (feature-öket). A szegmensek racionális reprezentációját az (2.2) egyenletben bemutatott optimalizációs probléma megoldásával állítottuk elő. Ebből egy ritka racionális reprezentációt gyártottunk, a komponensek 2.2 pontban bemutatott szelektálásával. Az így kapott ritka reprezentációk lokális bináris mintázatait felhasználva olyan jellemzőket nyertünk, melyek alapján az EEG iktális (rohamot ábrázoló) mintái jól elkülöníthetővé váltak a jel többi részétől. A jellemzőket egy logisztikus regressziós algoritmus felügyelt, páciens alapú tanításával választottuk szét roham/nem roham osztályokra. A módszert a PhysioNet CHB-MIT Scalp EEG adatbázis [GolAma00] egy 163 órás részhalmazán teszteltünk, melyen 70.4% szenzitivitást és 99.1% specificitást értünk el. A szegmensalapú osztályozás mellett megvizsgáltuk azt is, hogy milyen pontossággal tudjuk detektálni a rohamok

kezdetét. A tesztek során a szenzitivitás 91.13%, az óránkénti hamis riasztások aránya pedig átlagosan 0.35 volt. Ezzel számos state-of-the-art algoritmust, például wavelet, entrópia, idő-frekvencia spektrum alapú módszert sikerült megelőznünk.

4.2. Rendszeridentifikáció

A rendszeridentifikáció feladata egy szerkezetileg zárt rendszer viselkedésének (általában impulzus vagy frekvencia válaszának) modellezése. Ehhez a rendszer tet-szőleges bemeneti jelre adott válaszát tudjuk csak mérni, mely gyakran keveredik zajjal. A rendszer identifikációja során a bemeneti és a zajos kimeneti mérésekre kell matematikai modellt illeszteni (ld. 4.1). Attól függően, hogy a H rendszert és a v zajforrást hogyan modellezzük számos megközelítése van a problémának az idő- és frekvenciatérben [Lju99].



4.1. ábra. Egy H rendszer y kimenete az u bemenet és a v additív zaj függvényében.

Számos gyakorlati alkalmazásban fontos feladat, hogy a jel adott frekvenciájú komponenseit felismerjük, detektáljuk. Például telekommunikációs, illetve radaros alkalmazások esetén a vivőjel frekvenciájának vagy a közvetítő csatorna torzításának identifikálása is ilyen. A [GarKov20] munkában frekvencia identifikációval foglalkoztam, melyhez a 2. fejezetben bemutatott adaptív transzformációs módszereket használtam. A módszerrel egymáshoz közeli frekvenciájú komponensek is szétválaszthatók, mely alacsony jel-zaj arány (-15-0dB) mellett is pontosabb becslést tesz lehetővé, mint más state-of-the-art módszerek. A becsléshez tartozó (2.3) típusú problémát egy gradiens alapú optimalizációs módszerrel oldottuk meg, melyben leegyszerűsítettük a gradiens és a célfüggvény kiértékeléséhez szükséges $\Theta(\mathbf{a})^+$ pszeudoinverz kiszámítását. Így kapott implementációval [GarKov20] jelentősen csökkentettük az algoritmus futásidejét.

A szívzmokban terjedő elektromos ingerület időbeli terjedése is modellezhető rendszeridentifikációs módszerek segítségével. Ennek egy speciális esete a szívfrekvencia változásának (HRV) becslése. A szívütések gyakorisága könnyen számolható az EKG jelek egymás utáni R csúcsai közötti távolságok (RR intervallumok) reciprokával, mely jól tükrözi az ember általános egészségi állapotát. A tartósan fennálló

szívritmuszavarokat általában a szív szerkezeti károsodásai okozzák, az ideiglenes, rövid ideig tartó változások viszont nem feltétlenül jeleznek rendellenességet. Például a szívverés belégzés hatására kissé felgyorsul, kilégzés során pedig lelassul, ami teljesen normális. Ezek a külső behatások azonban torzíthatják a HRV becslését és megnehezítik az orvosi diagnózis felállítását. A nemzetközi irodalomban [ChoGut11; LenKir17] gyakran lineáris időinvariáns rendszeridentifikációs módszerek segítségével becsülik a HRV-t befolyásoló külső tényezőket (pl. légzést). Ezzel szemben, a [BöcKos19] munkában egy külső gerjesztéssel ellátott időben változó lineáris autoregresszív LPV-ARX modellt használtunk az identifikálandó H kardiorespirációs rendszer meghatározására. Ebben a megközelítésben a rendszert kívülről befolyásoló u jelet a légzési periódusoknak, az y kimenetet pedig az RR intervallumok hosszának választottuk. A becslések validálásához 7 ember bevonásával saját méréseket végeztünk a JKU-n. A mérések során arra kértük az önkénteseket, hogy egy előre megadott (0.3Hz, 0.4Hz, 0.5Hz) ütemben lélegezzenek, így megközelítőleg egy szakaszonként konstans u légzési gyakoriság mintázatot kaptunk. A tesztek során megmutattuk, hogy az így kapott LPV-ARX modell segítségével pontosabban lehet becsülni a HRV-t, mint a sima AR és ARX modellekkel.

A dinamikai rendszerek identifikációja állapottermodellek segítségével is megközelíthető, melyet az [LanKov18] munkában fotoakusztikus képrekonstrukcióhoz használtunk fel. A fotoakusztikus tomográfiában (PAT) a vizsgálandó anyagot, biológiai szövetet, stb. lézerrel világítják meg, melynek egy részét az anyag elnyeli ami hőenergiává alakul. A hőemelkedés hatására az anyag kitágul, megváltoztatva ezzel a körülvevő közeg nyomását. A megvilágítás periodikus ismétlésével periodikus nyomáskülönbség alakul ki a vizsgált objektum körül, amit ultrahang detektorokkal lehet mérni, a felszíni ultrahangos jelek tulajdonságaiból pedig következtetni lehet az azt gerjesztő közeg jellemzőire (pl. az elnyelődésre). A [LanKov18] cikkben a fizikai jelenséget leíró parciális differenciálegyenletek (PDE) diszkretizációjából levezettünk egy állapottermodellt, melynek segítségével algoritmust adtunk az inhomogén anyagok elnyelődési együtthatóinak becslésére. Numerikus szempontból több kihívással kellett szembenéznünk. Egyrészt a csillapított hullámterjedést leíró PDE diszkretizációjából származtatott állapottermodell számos paramétertől (időbeli/térbeli felbontás, relaxációs idő, stb.) függ, melyek értékei befolyásolják a modell matematikai tulajdonságait (pl. a stabilitást). Másrészt, a modelltől kapott $\mathbf{Ax} = \mathbf{b}$ típusú lineáris egyenletrendszer is rosszul kondicionált, ami miatt az becslés érzékeny a bemenő adatok perturbációjára, azaz a zajra. Ennek kiküszöbölésére a (3.5) egyenletben bemutatott regularizációs módszereket alkalmaztam, illetve meghatároztam a PDE azon paramétereit, melyekkel aszimptotikusan stabil, irányítható és megfigyelhető állapotter reprezentációkat kapunk. A módszer nagy előnye, hogy a modellen keresztül optimalizálni lehet a külső gerjesztéshez használt u lézer modulációs paramétereit is, amivel javítható az elnyelődési együtthatók becslése.

Saját publikációk (2016–2021)

Folyóiratcikkek

- [BöcKos19] C. Böck, K. Kostoglou, P. Kovács, M. Huemer és J. Meier. “A Linear Parameter Varying ARX Model for Describing Biomedical Signal Couplings”. *Computer Aided Systems Theory–EUROCAST 2011: Part II*. Szerk. R. Moreno-Díaz et al. 12014. köt. LNCS. **Q3**. Heidelberg, Germany: Springer-Verlag Berlin, 2019, 339–346. old. DOI: 10.1007/978-3-030-45096-0_42.
- [BöcKov21] C. Böck, P. Kovács, P. Laguna, J. Meier és M. Huemer. “ECG Beat Representation and Delineation by means of Variable Projection”. *IEEE Transactions on Biomedical Engineering* (2021). **Q1**, 1–12. old. DOI: 10.1109/TBME.2021.3058781.
- [DózBog19] T. Dózsa, G. Bognár és P. Kovács. “Ensemble Learning for Heartbeat Classification Using Adaptive Orthogonal Transformations”. *Computer Aided Systems Theory–EUROCAST 2019: Part II*. Szerk. R. Moreno-Díaz et al. 12014. köt. LNCS. **Q3**. Heidelberg, Germany: Springer-Verlag Berlin, 2019, 355–363. old. DOI: 10.1007/978-3-030-45096-0_44.
- [Kov17b] P. Kovács. “Rational Variable Projection Methods in ECG Signal Processing”. *Computer Aided Systems Theory–EUROCAST 2017: Part II*. Szerk. R. Moreno-Díaz et al. 10672. köt. LNCS. **Q2**. Heidelberg, Germany: Springer-Verlag Berlin, 2017, 196–203. old. DOI: 10.1007/978-3-319-74727-9_23.
- [KovBog21b] P. Kovács, G. Bognár, C. Huber és M. Huemer. “VPNet: Variable Projection Networks”. *International Journal of Neural Systems (IJNS)* (2021). **D1**, (under review), 1–17. old.
- [KovFek19] P. Kovács és M. Fekete A. “Nonlinear least-squares spline fitting with variable knots”. *Applied Mathematics and Computation* 354 (2019). **Q1**, 490–501. old. DOI: 10.1016/j.amc.2019.02.051.
- [KovFri20] P. Kovács, S. Fridli és F. Schipp. “Generalized Rational Variable Projection with Application in ECG Compression”. *IEEE Transactions on Signal Processing* 68 (2020). **D1**, 478–492. old. DOI: 10.1109/TSP.2019.2961234.
- [KovLeh20b] P. Kovács, B. Lehner, G. Thummerer, G. Mayr, P. Burgholzer és M. Huemer. “Deep learning approaches for thermographic imaging”. *Journal of Applied Physics* 128.15 (2020). **Q2**, 155103-1–16. old. DOI: 10.1063/5.0020404.

- [LanKov18] O. Lang, P. Kovács, C. Motz, M. Huemer, T. Berer és P. Burgholzer. “A linear state space model for photoacoustic imaging in an acoustic attenuating media”. *Inverse Problems* 35 (2018). **Q1**, 1–29. old. DOI: 10.1088/1361-6420/aaea2c.
- [LehGal21] B. Lehner és tsai. “Uncertainty Estimation for Deep Learning-based Thermographic Imaging”. *Sensors & Transducers* 249.2 (2021). ISSN: 2306-8515, 25–35. old.
- [SamKov17] K. Samiee, P. Kovács és M. Gabbouj. “Epileptic seizure detection in long-term EEG records using sparse rational decomposition and local Gabor binary patterns feature extraction”. *Knowledge-Based Systems* 118 (2017). **D1**, 228–240. old. DOI: 10.1016/j.knsys.2016.11.023.

Konferenciák

- [BogFri19] G. Bognár, S. Fridli, P. Kovács és F. Schipp. “Adaptive Rational Transformations in Biomedical Signal Processing”. *Progress in Industrial Mathematics at ECMI*. Springer-Verlag Berlin, 2019, 239–247. old. DOI: 10.1007/978-3-030-27550-1_30.
- [GarKov20] Y. E. Garcia, P. Kovács és M. Huemer. “Variable Projection for Multiple Frequency Estimation”. *Proceedings of the 45th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, 4811–4815. old. DOI: 10.1109/ICASSP40776.2020.9052907.
- [Kov18] P. Kovács. “Sparsity problem involving rational basis functions”. *Proceedings of the 21st IEEE Statistical Signal Processing Workshop (SSP)*. IEEE, 2018, 618–622. old. DOI: 10.1109/SSP.2018.8450725.
- [KovBöc17] P. Kovács, C. Böck, J. Meier és M. Huemer. “ECG Segmentation Using Adaptive Hermite Functions”. *Proceedings of the 51st Annual Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2017, 1476–1480. old. DOI: 10.1109/ACSSC.2017.8335601.
- [KovBöc19] P. Kovács, C. Böck, T. Dózsa, J. Meier és M. Huemer. “Waveform Modeling by Adaptive Weighted Hermite Functions”. *Proceedings of the 44th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, 1080–1084. old. DOI: 10.1109/ICASSP.2019.8683296.
- [KovLeh20a] P. Kovács, B. Lehner, G. Thummerer, G. Mayr, P. Burgholzer és M. Huemer. “A Hybrid Approach for Thermographic Imaging With Deep Learning”. *Proceedings of the 45th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, 4277–4281. old. DOI: 10.1109/ICASSP40776.2020.9053411.

Szoftverek

- [GarKov20] Y. E. Garcia és P. Kovács. *Variable Projection for Multiple Frequency Estimation*. Available: <https://codeocean.com/capsule/5263510/tree/v1>. 2020.
- [Kov17a] P. Kovács. *Generalized Rational Variable Projection with Application in ECG Compression*. <https://codeocean.com/capsule/8330505/tree/v1>. 2017.
- [KovBog21a] P. Kovács, G. Bognár és C. Huber. *VPNet: Variable Projection Networks*. Available: <https://git.silicon-austria.com/pub/sparseestimation/vpnet>. 2021.
- [KovBöc18] P. Kovács és C. Böck. *Waveform modeling by adaptive weighted Hermite functions*. <https://codeocean.com/capsule/6391724/tree/v1>. 2018.
- [KovLeh20c] P. Kovács, B. Lehner, G. Thummerer, G. Mayr, P. Burgholzer és M. Huemer. *ThermUnet - Deep Learning Approaches for Thermographic Imaging*. Available: <https://git.silicon-austria.com/pub/confine/ThermUNet>. 2020.

Referenciák

- [AmoKol17] B. Amos és J. Z. Kolter. “OptNet: differentiable optimization as a layer in neural networks”. *34th International Conference on Machine Learning (ICML 2017)*. PMLR, 2017, 136–145. old.
- [BurTho17] P. Burgholzer, M. Thor, J. Gruber és G. Mayr. “Three-dimensional thermographic imaging using a virtual wave concept”. *Journal of Applied Physics* 121.10 (2017), 105102: 1–11.
- [ChaMen18] B. Chang, L. Meng, E. Haber, L. Ruthotto, D. Begert és E. Holtam. “Reversible Architectures for Arbitrarily Deep Residual Neural Networks”. *The 32th AAAI Conference on Artificial Intelligence (AAAI-18)*. AAAI, 2018, 2811–2818. old.
- [CheDon01] S. S. Chen, D. L. Donoho és M. A. Saunders. “Atomic Decomposition by Basis Pursuit”. *SIAM Review* 43.1 (2001), 129–159. old.
- [CheRub18] R. T. Q. Chen, Y. Rubanova, J. Bettencourt és D. Duvenaud. “Neural Ordinary Differential Equations”. *32nd Conference on Neural Information Processing Systems (NeurIPS 2018)*. NIPS, 2018, 6572–6583. old.
- [ChoGut11] J. Choi és R. Gutierrez-Osuna. “Removal of Respiratory Influences From Heart Rate Variability in Stress Monitoring”. *IEEE Sensors Journal* 11.11 (2011), 2649–2656. old.

- [Cyb89] G. Cybenko. “Approximation by superpositions of a sigmoidal function”. *Mathematics of Control, Signals and Systems* 2 (1989), 303–314. old.
- [DózKov15] T. Dózsa és P. Kovács. “ECG Signal Compression Using Adaptive Hermite Functions”. *Advances in Intelligent Systems and Computing* 399 (8 2015).
- [FriLóc12] S. Fridli, L. Lócsi és F. Schipp. “Rational function system in ECG processing”. *Computer Aided Systems Theory–EUROCAST 2011: Part I*. Szerk. R. Moreno-Díaz et al. 6927. köt. LNCS. Heidelberg, Germany: Springer-Verlag Berlin, 2012, 88–95. old.
- [GolAma00] A. L. Goldberger és tsai. “PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals”. *Circulation* 101.23 (2000), 215–220. old.
- [GolPer03] G. H. Golub és V. Pereyra. “Separable nonlinear least squares: The variable projection method and its applications”. *Inverse problems* 19.2 (2003), R1–R26.
- [GolPer73] G. H. Golub és V. Pereyra. “The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate”. *SIAM Journal on Numerical Analysis (SINUM)* 10 (1973), 413–432. old.
- [Han92] P. C. Hansen. “Analysis of Discrete Ill-Posed Problems by Means of the L-Curve”. *SIAM Review* 34.4 (1992), 561–580. old.
- [HerRou14] J. R. Hershey, J. Le Roux és F. Weninger. *Deep Unfolding: Model-Based Inspiration of Novel Deep Architectures*. arXiv: 1409.2574. 2014. arXiv: 1409.2574 [cs.LG].
- [HeZha16] K. He, X. Zhang, S. Ren és J. Sun. “Deep Residual Learning for Image Recognition”. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2016, 770–778. old.
- [Hor91] K. Hornik. “Approximation capabilities of multilayer feedforward networks”. *Neural Networks* 4.2 (1991), 251–257. old.
- [Jup78] D. L. B. Jupp. “Approximation to data by splines with free knots”. *SIAM Journal on Numerical Analysis* 15.2 (1978), 328–343. old.
- [KarGab97] M. Karczewicz és M. Gabbouj. “ECG data compression by spline approximation”. *Signal Processing* 59 (1997), 43–59. old.
- [Kay93] S. M. Kay. *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. New Jersey, USA: Prentice Hall PTR, 1993.
- [KenEbe95] J. Kennedy és R. C. Eberhart. “Particle swarm optimization”. *Proceedings of IEEE International Conference on Neural Networks*. 4. köt. 1995, 1942–1948. old.

- [KidLyo20] P. Kidger és T. Lyons. “Universal Approximation with Deep Narrow Networks”. *Proceedings of Thirty Third Conference on Learning Theory (PMLR)* 125 (2020), 2306–2327. old.
- [Law13] M. Lawrińczuk. “Practical nonlinear predictive control algorithms for neural Wiener models”. *Journal of Process Control* 23.5 (2013), 696–714. old.
- [LenKir17] G. Lenis, M. Kircher, J. Lázaro, R. Bailón és E. Gil. “Separating the effect of respiration on the heart rate variability using Granger’s causality and linear filtering”. *Biomedical Signal Processing and Control* 31 (2017), 272–287. old.
- [Lju99] L. Ljung. *System identification: Theory for the user*. 2nd. New Jersey, USA: Prentice Hall, 1999.
- [MalZha93] S. G. Mallat és Z. Zhang. “Matching pursuits with time-frequency dictionaries”. *IEEE Trans. Signal Process.* 41.12 (1993), 3397–3415. old.
- [Nat52] I. P. Natanszon. *Konstruktív függvénytan*. Szerk. B. Szőkefalvi-Nagy. Ford. A. Rényi. Budapest, Hungary: Akadémiai Kiadó, 1952.
- [Neb98] F. Nebeker. *The IEEE Signal Processing Society: Fifty Years of Service, 1948-1998*. IEEE History Center, 1998.
- [PapSch18] M. Pap és F. Schipp. “Quaternionic Blaschke Group”. *Mathematics* 7.1 (2018), 1–12. old.
- [RonFis15] O. Ronneberger, P. Fischer és T. Brox. “U-net: Convolutional networks for biomedical image segmentation”. *International Conference on Medical image computing and computer-assisted intervention*. 2015, 234–241. old.
- [Sze67] G. Szegő. *Orthogonal polynomials*. 3rd. New York, USA: AMS Colloquium Publications, 1967.
- [Szi07] L. Szili. *Funkcionálanalízis – a jelfeldolgozás és a szimuláció matematikai alapjai*. Budapest, HU: ELTE IK, Kari Digitális Könyvtár, 2007.
- [TanZha19] C. Tan, L. Zhang és H. Wu. “A Novel Blaschke Unwinding Adaptive Fourier Decomposition based Signal Compression Algorithm with Application on ECG Signals”. *IEEE Journal of Biomedical and Health Informatics* 23 (2 2019).
- [Tib09] R. Tibshirani. “Regression shrinkage and selection via the lasso”. *Journal of the Royal Statistical Society. Series B (Methodological)* 58.1 (2009), 267–288. old.
- [Tro04a] J. A. Tropp. “Greed is good: algorithmic results for sparse approximation”. *IEEE Trans. Inf. Theory* 50.10 (2004), 2231–2242. old.

- [Tro04b] J. A. Tropp. “Just Relax: Convex Programming Methods for Identifying Sparse Signals in Noise”. *IEEE Trans. Inf. Theory* 52.3 (2004), 1030–1051. old.
- [UyaÜlk17] K. Uyar és E. Ülker. “B-Spline Curve Fitting With Invasive Weed Optimization”. *Applied Mathematical Modelling* 52 (2017).
- [WilNoë19] M. J. Willemink és P. B. Noël. “The evolution of image reconstruction for CT—from filtered back projection to artificial intelligence”. *European Radiology* 29.5 (2019), 2185–2195. old.
- [YosHar03] F. Yoshimoto, T. Harada és Y. Yoshimoto. “Data fitting with a spline using a real-coded genetic algorithm”. *Computer-Aided Design* 35 (8 2003).
- [YuaChe13] Y. Yuan, N. Chen és S. Zhou. “Adaptive B-spline knot selection using multi-resolution basis set”. *IIE (Institute of Industrial Engineers) Transactions* 45 (2013).
- [YuPen14] H. Yu, J. Peng és Y. Tang. “Identification of Nonlinear Dynamic Systems Using Hammerstein-Type Neural Network”. *Mathematical Problems in Engineering* 2014 (2014).