

České vysoké učení technické v Praze
Fakulta jaderná a fyzikálně inženýrská

Výzkumný úkol

Emil Doležal

České vysoké učení technické v Praze
Fakulta jaderná a fyzikálně inženýrská

Analýza EEG signálu pomocí teorie náhodných matic

Katedra fyziky

Akademický rok: 2005/2006

Školitel: Prof. RNDr. Petr Šeba, DrSc.

Rád bych poděkoval prof. RNDr. Petru Šebovi, DrSc. za poskytnuté konzultace, cenné rady a připomínky.

Obsah

Úvod	7
1 Spektrální analýza barevného šumu	9
1.1 Konstrukce korelačních matic	9
1.2 Marchenko-Pasturova předpověď rozdělení vlastních čísel	10
1.3 Frekvenční analýza časových řad	10
1.4 Spektrální hustota korelačních matic obecného barevného šumu	11
1.5 Zobecněný bílý šum	13
1.6 Speciální případ barevného šumu	15
2 Analýza lidského elektroencefalogramu	19
2.1 Měření EEG	19
2.2 Frekvenční charakteristika	20
2.3 Hustota vlastních čísel korelačních matic	21

Úvod

Mozek je asi nejsložitější a nejdůležitější orgán v lidském těle. Svou činností řídí všechny ostatní části organismu. Přes velké úsilí doktorů a vědců z jiných oborů, není jeho popis zdaleka úplný. Procesy probíhající v mozku jsou velmi komplikované. Na první pohled se zdají být dokonce chaotické a neuspořádané.

Mozek patří mezi komplexní systémy. Jejich charakteristickým rysem je složitost, díky které vypadají z vnějšku chaoticky, ale ve skutečnosti nejsou. Detailní znalost interakce mezi stavebními kameny komplexních systémů nezaručuje pochopení chování celého objektu. Teorie náhodných matic poskytuje nástroje, jak takové systémy studovat a klasifikovat.

Elektromagnetická aktivita mozku se měří elektrodami, rozmístěnými na povrchu lebky. Hlavním zdrojem aktivity jsou nejvyšší vrstvy kortexu (koncového mozku), obsahující neurony. Vznikající signály se rozdělují do několika typů podle umístění zdroje, intenzity a frekvenční charakteristiky.

Elektrické vlastnosti povrchu a obsahu lebky umožňují šíření jen určitým vlnám. Součinností neuronů vzniká střídavé napětí s frekvencí v rozsahu 4-50 Hz. Ve výsledku vypadá signál naměřený elektrodami jako barevný šum.

V první kapitole se budeme zabývat spektrálními vlastnostmi barevného šumu, které ve druhé kapitole použijeme k analýze lidského elektroencefalogramu (EEG).

Kapitola 1

Spektrální analýza barevného šumu

Prostor kolem nás je vyplněn vlněním rozličné povahy. Existuje mnoho cest, vedoucích k jeho vytvoření. Jmenujme alespoň některé. Chvěním materiálů vznikají vlny zvukové (princip hudebních nástrojů). Při kmitání molekul nebo atomů v krystalické mřížce se uvolňuje tepelné záření. Zrychlený pohyb nabitých částic je zdrojem vlnění elektromagnetického. Bez ohledu na typ zdroje můžeme každé vlně přisoudit amplitudu, frekvenci a rychlost šíření.

Vzhledem k různé povaze vlnění existuje také pestrá paleta detektorů, konstruovaných k měření konkrétních druhů vln. Detektory periodicky měří hodnotu jisté fyzikální veličiny (např. napětí), jejich výstupem jsou proto časové řady.

Střetnou-li se vlny na jednom místě, dochází k jejich skládání (superposici). Umístíme-li do tohoto místa detektor, nedostaneme již informaci o jednotlivých vlnách, ale pouze o jejich superposici. Frekvenčním rozkladem časových řad, zaznamenaných detektorem, můžeme alespoň určit zastoupení signálu o daném kmitočtu.

Předpokládejme nadále, že signál je náhodný, gaussovsky rozdělený, se střední hodnotou rovnou nule. Jsou-li v rozkladu zastoupeny všechny frekvence a přispívají-li k celkovému výkonu přibližně stejně, mluvíme v souvislosti s přicházejícím vlněním jako o bílém šumu. Nepřispívají-li všechny frekvence k celkovému výkonu rovnoměrně, budeme vlnění nazývat barevný šum. Přesná definice bude zmíněna později.

Data získaná z jednoho detektoru většinou nejsou dostatečná k analýze měřeného signálu. Z tohoto důvodu se do zkoumané oblasti umísťuje celá sada stejných detektorů. Naměřené časové řady je možné rozdělit na okénka předepsané délky. Z údajů v okénkách následně vytvořit korelační matice a zkoumat jejich vlastní čísla. V případě, že frekvenční analýza dat ze všech detektorů ukazuje na bílý šum, předpovídá teorie vybudovaná Pasturem a Marchenkem [1], za jistých okolností, tvar rozdělovací funkce vlastních čísel. Cílem této kapitoly je rozšířit předpověď o tvaru rozdělovací funkce bílého šumu na barevný šum.

1.1 Konstrukce korelačních matic

Uvažujme systém M detektorů poskytující M časových řad $\{x^i\}_{i=1}^M$. Dále předpokládejme, že všechna měřidla zaznamenala N hodnot.

Prvním krokem při zpracování dat je rozdělení časových řad na okénka. Tento krok je velmi důležitý, protože správná volba počtu měření n v jednom okénku ovlivňuje celou další analýzu. Z následujících úvah bude zřejmé, že n musí být alespoň stejně velké jako M , jinak bychom s jistotou dostali singulární korelační matice. Naopak, je-li n srovnatelné s N , nemůžeme z výsledků dosažených pro jednotlivá okénka říci, ke které události se vztahují (v jednom okénku může proběhnout více událostí).

Časovou řadu x^i je zvykem přeškálovat tak, aby platilo

$$\sum_{j=n_1}^{n_2} x_j^i = 0, \quad \sum_{j=n_1}^{n_2} (x_j^i)^2 = 1, \quad (1.1)$$

kde n_1, n_2 určují počátek a délku okénka $n_2 - n_1 = n$. Této úpravě řad se říká whitening - vybělení.

Definujeme prvek $c_{i,m}$ korelační matice vztahem

$$c_{i,j} = \sum_{k=n_1}^{n_2} x_k^i x_k^j. \quad (1.2)$$

Označíme-li symbolem x matici $M \times n$ obsahující data z časových řad v řádcích, můžeme předchozí definici zapsat přehledně $c = xx^T$.

1.2 Marchenko-Pasturova předpověď rozdělení vlastních čísel

Nechť M je i nadále počet detektorů, které používáme při analýze. Zvolíme-li délku měření N nepřilíši velkou a okénko dostatečně široké vzhledem k celému souboru, poskytují korelační matice pro jednotlivé vzorky dobrou aproximaci korelační matice pro celý soubor. Nicméně, necháme-li N růst do nekonečna a šířka okénka n zůstane konečná, zjistíme, že výše zmiňovaný vztah již nebude platit.

Za předpokladu, že prvky matice x jsou nezávislé a identicky rozdělené kolem nuly se stejnou variací σ^2 , předpovídá Marchenko-Pasturovův teorém (viz [1]) v limitním případě $M \rightarrow \infty$, $\frac{M}{n} \rightarrow c$ pro korelační matici xx^T příslušející k jednomu okénku hustotu vlastních čísel ve tvaru

$$\varrho(\lambda) = \begin{cases} \frac{1}{2\pi\lambda c\sigma^2} \sqrt{(b-\lambda)(\lambda-a)} & a \leq \lambda \leq b \\ 0 & \lambda < a \vee \lambda > b \end{cases}, \quad (1.3)$$

kde $a = \sigma^2(1 - \sqrt{c})^2$, $b = \sigma^2(1 + \sqrt{c})^2$ a $0 < c \leq 1$. Při volbě $c > 1$ se v hustotě $\varrho(\lambda)$ dodatečně objeví Diracova delta funkce v bodě 0 o velikosti $1 - \frac{1}{c}$. Existence delta funkce v hustotě vlastních čísel je zřejmá (viz odstavec 1.1), uvážíme-li fakt, že v tomto případě dostáváme singulární korelační matice.

V dalších odstavcích zjistíme, že rozdělení (1.3) je do značné míry univerzální. Ukážeme, že je speciálním případem obecnějšího rozdělení vlastních čísel, závislého na frekvenčním rozkladu časových řad, z nichž je vytvořena příslušná korelační matice.

1.3 Frekvenční analýza časových řad

Důležitou charakteristikou měřicího přístroje, periodicky zaznamenávajícího určitou fyzikální veličinu, je vzorkovací frekvence f_v . Udává, jak často přístroj ukládá naměřenou hodnotu. Nepřímo také ovlivňuje rozsah frekvencí pro frekvenční analýzu.

Časovou řadu si můžeme představit jako prvek vektorového prostoru \mathbb{R}^N . Podobně data v jednotlivých okénkách této řady jako vektory v prostoru \mathbb{R}^n . Zavedená interpretace umožňuje podrobit časovou řadu diskrétní Fourierově transformaci.

Diskrétní Fourierova transformace (DFT) je definována na komplexních prostorech formulemi (složky vektorů číslujeme od 0)

$$X_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{-i\omega_k t_j}, \quad x_j = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X_k e^{i\omega_j t_k} \quad (1.4)$$

kde úhlová frekvence ω_k a čas t_j jsou určeny vzorkovací frekvencí

$$\omega_k = k \frac{2\pi}{N} f_v = k \frac{2\pi}{TN}, \quad t_j = \frac{j}{f_v} = jT, \quad f_v = \frac{1}{T}, \quad (1.5)$$

T je časový interval mezi dvěma měřeními.

Transformační vztahy (1.4) je možné zapsat přehledně v operátorovém formalismu. Nechť U označuje operátor definovaný prvním výrazem v (1.4). Tento operátor je unitární, zachovává tedy standardní skalární součin \langle, \rangle . Inverse U^{-1} je vyjádřena druhou formulí v (1.4). Zastavme se u několika vlastností operátoru diskrétní Fourierovy transformace U .

Reálnost vektoru x umožňuje přepis definice inverzního operátoru U^{-1} na tvar

$$x_j = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \operatorname{Re}(X_k) \sin \omega_j t_k + \operatorname{Im}(X_k) \cos \omega_j t_k, \quad (1.6)$$

ze kterého je vidět, že reálné a imaginární části složek vektoru X_k představují v jistém smyslu amplitudy (nemusí však být kladné).

Definici korelační matice (1.2) můžeme pomocí skalárního součinu zapsat jako rovnost

$$c_{l,m} = \langle x^l, x^m \rangle_{\mathbb{R}^N} = \langle X^l, X^m \rangle_{\mathbb{C}^N}, \quad (1.7)$$

kde indexy \mathbb{R}^N a \mathbb{C}^N určují prostor, na kterém skalární součin počítáme.

Operátor U , zúžený na reálný prostor \mathbb{R}^N , zobrazuje vektor $x \in \mathbb{R}^N$ na vektor $X \in \mathbb{C}^N$, který je obecně komplexní a jeho složky nejsou nezávislé. Přesněji: platí vztah

$$X_i = \overline{X_{N-i}}. \quad (1.8)$$

V analýze signálu se používá charakteristika nazývaná výkonové (power) spektrum. Jedná se o graf závislosti $|X_k|^2$ na odpovídající frekvenci ω_k . Určuje příspěvky jednotlivých frekvencí k celkovému výkonu, který je přímo úměrný kvadrátu amplitud. Díky platnosti vztahu (1.8) je zbytečné kreslit power spektrum pro všechny frekvence ω_k , obvykle se použije prvních $\left[\frac{N}{2}\right]$, kde symbol $[a]$ znamená celou část z a .

Některé vlastnosti vektoru X jsou pro naši další analýzu nežádoucí, proto se pokusíme definovat jiný lineární operátor \tilde{U} , zobrazující $\mathbb{R}^N \rightarrow \mathbb{R}^N$ tak, aby složky obrazu \tilde{X} vektoru x stále určovaly frekvenční rozklad (1.6) a aby skalární součin zůstal zachován. Definujeme působení operátoru \tilde{U} na vektor x pomocí vektoru $X = Ux$ vztahem

$$\tilde{X}_i = \sqrt{2} \cdot \begin{cases} \operatorname{Re}(X_i) & 0 < i < \left[\frac{N}{2}\right] \\ \operatorname{Im}\left(X_{i-\left[\frac{N}{2}\right]+1}\right) & \left[\frac{N}{2}\right] < i \leq N-1 \end{cases}. \quad (1.9)$$

Zbývající dva prvky definujeme samostatně, protože jejich význam a vlastnosti se od ostatních liší. První prvek X_0 je vždy reálný a je definován součtem řady $\{x_i\}$. Pro sudé N je prvek $X_{\frac{N}{2}}$ také reálný a označuje se "Nyquist component". Obě posledně zmíněné složky vektoru X jsou pro frekvenční rozklad v jistém smyslu zbytečné, jelikož nehrají roli amplitudy vlny o jisté frekvenci jako ostatní složky. Definujeme $\tilde{X}_0 = X_0$ a v závislosti na N , buď $\tilde{X}_{\frac{N}{2}} = X_{\frac{N}{2}}$ pro sudé N , nebo $\tilde{X}_{\frac{N-1}{2}} = \sqrt{2}X_{\frac{N-1}{2}}$, je-li N liché.

Snadno ověříme, že takto definovaný operátor má všechny požadované vlastnosti. Podobně jako v odstavci 1.1 můžeme definici korelační matice přehledně zapsat rovností

$$c = xx^T = XX^* = \tilde{X}\tilde{X}^T, \quad (1.10)$$

kde x je matice $M \times n$ a matice X, \tilde{X} jsou vytvořeny analogickým postupem jako x . Zobrazení-li závislost \tilde{X}_k^2 na frekvenci ω_k dostaneme charakteristiku, kterou budeme v dalších odstavcích nazývat modifikované power spektrum. Jeho význam je podobný jako u power spektra, liší se v tom, že modifikované power spektrum rozlišuje mezi sudou a lichou částí vlny $X_k e^{i\omega_j t_k}$.

1.4 Spektrální hustota korelačních matic obecného barevného šumu

V předchozím odstavci jsme popsaly nástroje, kterými je možné odlišit barevný šum od bílého šumu. Následující definice jsou založeny na tvaru power spektra.

1.4.1 Definice (Gaussovský bílý šum). Mějme náhodnou časovou řadu $\{x_j\}_{j=0}^{n-1}$, jejíž elementy jsou nezávislé, gaussovsky rozdělené kolem nuly. Je-li power spektrum příslušné k řadě $\{x_j\}_{j=0}^{n-1}$ rovnoměrné, nazýváme odpovídající signál bílým šumem.

1.4.2 Definice (Gaussovský barevný šum). Mějme náhodnou časovou řadu $\{x_j\}_{j=0}^{n-1}$, jejíž elementy jsou nezávislé, gaussovsky rozdělené kolem nuly. Není-li power spektrum příslušné k řadě $\{x_j\}_{j=0}^{n-1}$ rovnoměrné, nazýváme odpovídající signál barevným šumem.

V důsledku nulovosti středních hodnot složek řady $\{x_j\}_{j=0}^{n-1}$ jsou také střední hodnoty složek rozkladu $\{\tilde{X}_k\}_{k=0}^{n-1}$ časové řady $\{x_j\}_{j=0}^{n-1}$ nulové (viz. definiční vztahy 1.4 a 1.9). Dále pro členy diskrétního furierova rozkladu $\{X_k\}_{k=0}^{n-1}$ platí

$$\langle (\operatorname{Re}(X_k))^2 \rangle = \langle (\operatorname{Im}(X_k))^2 \rangle, \quad \text{pro } 1 \leq k \leq \left\lfloor \frac{N-1}{2} \right\rfloor, \quad (1.11)$$

neboli

$$\langle (\tilde{X}_k)^2 \rangle = \left\langle \left(\tilde{X}_{\left\lfloor \frac{N-1}{2} \right\rfloor + k} \right)^2 \right\rangle, \quad \text{pro } 1 \leq k \leq \left\lfloor \frac{N-1}{2} \right\rfloor. \quad (1.12)$$

Analogické definice k 1.4.1 a 1.4.2 bychom mohli formulovat na základě modifikovaného výkonového spektra, které by bylo symetrické ve smyslu rovnosti 1.12.

Z odstavce 1.2 víme, jak bude vypadat hustota vlastních čísel $\varrho(\lambda)$ korelačních matic xx^T , obsahují-li matice x v řádcích bílý šum. K určení spektrální hustoty korelačních matic odpovídajících barevnému šumu použijeme větu formulovanou Girkem (viz [2])

1.4.3 Věta. Buď A náhodná matice $n \times [cn]$ s nezávislými elementy, které mají nulovou střední hodnotu a splňují podmínku

$$n\operatorname{Var}(A_{ij}) < B \quad (1.13)$$

pro $B < \infty$. Předpokládejme, že funkce $v_n : \langle 0, 1 \rangle \times \langle 0, c \rangle \rightarrow \mathbb{R}$ definované jako

$$v_n(x, y) = n\operatorname{Var}(A_{ij}), \quad \frac{i}{n} \leq x \leq \frac{i+1}{n}, \quad \frac{j}{n} \leq y \leq \frac{j+1}{n} \quad (1.14)$$

stejněměrně konvergují k limitní funkci $v(x, y)$. Potom limitní rozdělovací funkce vlastních čísel $\varrho(\lambda)$ korelační matice AA^T existuje a splňuje rovnici:

$$\int_0^\infty \frac{\varrho(\lambda)d\lambda}{1+t\lambda} = \int_0^1 u(x, t)dx, \quad (1.15)$$

kde funkce $u(x, t)$ je určena rovnicí:

$$u(x, t) = \frac{1}{1+t \int_0^c \frac{v(x, y)dy}{1+t \int_0^1 u(z, t)v(z, y)dz}}. \quad (1.16)$$

Řešení rovnice (1.16) existuje a je jednoznačné na třídě funkcí $u(x, t) \geq 0$, analytických v t a spojitých v $x \in \langle 0, 1 \rangle$.

Nechť matice \tilde{X} je obrazem matice x při zobrazení \tilde{U} (přesněji: řádky matice \tilde{X} jsou obrazem řádků matice x , viz odstavce 1.3). Odpovídají-li řádky matice x , o rozměrech $M \times n$, barevnému šumu, zvolíme \tilde{X} jako matici A ve Větě 1.4.3.

Ihned vidíme, že $c = \left\lfloor \frac{n}{M} \right\rfloor$. Splnění podmínky (1.13) nebudeme nyní řešit, vyplyne totiž z vlastností konkrétních posloupností $\{v_n\}$.

Modifikované výkonové spektrum, určující variaci složek v řádku matice \tilde{X} , je pro všechny řádky stejné. Funkce $v(x, y)$ díky tomu nebude záviset na prvním argumentu, tedy:

$$v(x, y) = v(y). \quad (1.17)$$

Využijeme-li vztahu (1.17), zjistíme na základě definice (1.16), že ani funkce $u(x, t)$ nezávisí na prvním argumentu.

Nakonec se zastavme u rovnic (1.15) a (1.16). Jejich tvar se také výrazně zjednoduší. V každé z nich zmizí jeden integrál a dostanou tvar

$$\int_0^{\infty} \frac{\varrho(\lambda) d\lambda}{1 + t\lambda} = u(t) \quad (1.18)$$

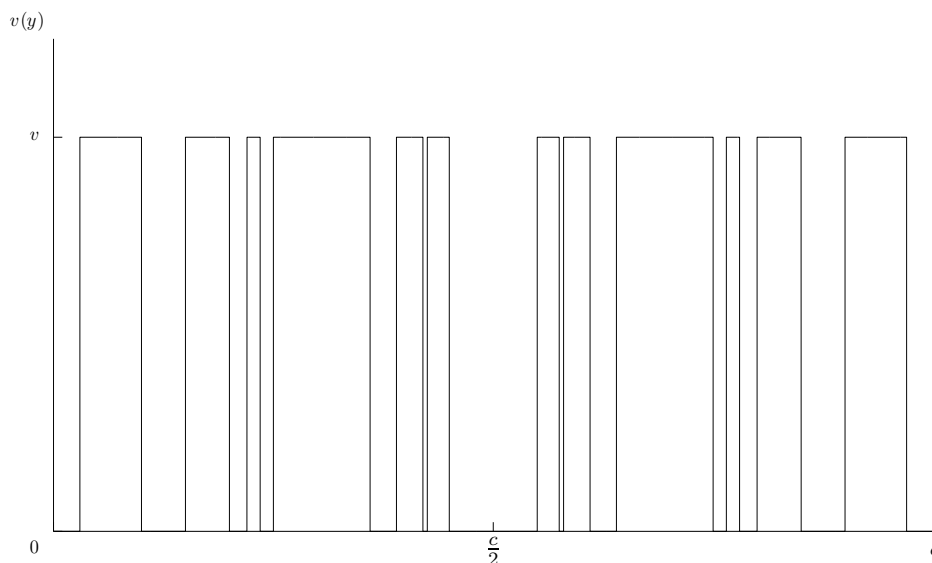
a

$$u(t) = \frac{1}{1 + t \int_0^c \frac{v(y) dy}{1 + tu(t)v(y)}}. \quad (1.19)$$

Principiálně jsme našli hustotu vlastních čísel libovolných korelačních matic $\tilde{X}\tilde{X}^T$, odpovídajících barevnému šumu. Řešení rovnic (1.18) a (1.19) však není triviální. V dalších odstavcích se pokusíme najít předpis pro hustotu $\varrho(\lambda)$, v případě, že funkce $v(y)$ je schodovitá (nespojité funkce tvořená konstantními bloky různé výšky).

1.5 Zobecněný bílý šum

Představme si situaci, kdy do jednoho místa přichází signál z několika zdrojů. Každý zdroj produkuje ve svém frekvenčním pásu změť vln náhodných fází. Společná je pro všechny vysílače pouze průměrná hodnota výkonu, přenášeného jednotlivými vlnami. Měříme-li přicházející signál M detektory, dostaneme časové řady $\{x_j^i\}_{i=1, j=0}^{i=M, j=n-1}$. Zobrazíme-li tyto řady diskretní fourierovou transformací a následně zkonstruujeme výkonové spektrum, bude příslušná funkce $v(y)$ (viz 1.17) vypadat podobně jako na obrázku 1.1 (stejně pro všech M řad!).



Obrázek 1.1: Graf funkce $v(y)$ pro zobecněný bílý šum

Z M časových řad $\{x_j^i\}_{i=1, j=0}^{i=M, j=n-1}$ umístěných do řádků vytvoříme, stejně jako v odstavci 1.1, matici x o rozměrech $M \times n$.

Pro určení hustoty vlastních čísel $\varrho(\lambda)$ korelačních matic xx^T jsou podstatné pouze dva údaje o tvaru funkce $v(y)$ diskutovaného typu. Součet délek intervalů d , na kterých $v(y)$ nabývá nenulové hodnoty, a funkční hodnota v na těchto intervalech. Není složité se přesvědčit, že řešení rovnice (1.19) nezávisí na umístění intervalů, na nichž je $v(y)$ konstantně rovna v .

Dosažením funkce $v(y)$, podobného tvaru jako je na obrázku 1.1, do výrazu (1.19), nalezneme předpis pro funkci $u(t)$

$$u(t) = \frac{-[1 + tv(d-1)] \pm \sqrt{[1 + tv(d-1)]^2 + 4tv}}{2tv}, \quad (1.20)$$

který je prozatím nejednoznačný. Věta 1.4.3 předpokládá řešení (1.19) jednoznačně na třídě analytických funkcí, pro než platí $u(t) \geq 0$. Znaménko před odmocninou ve vzorci (1.20) proto bude kladné.

Integrální rovnici (1.18) převedeme na tvar

$$\int_0^\infty \frac{\varrho(\lambda)d\lambda}{y+\lambda} = \frac{u\left(\frac{1}{y}\right)}{y} = \tilde{u}(y), \quad (1.21)$$

který odpovídá Stieltjesově transformaci (viz [3]).

Operátor zobecněné Stieltjesovy transformace je definován vztahem

$$G(z) = \int_0^\infty \frac{F(y)dy}{(y+z)^p}, \quad |\arg z| < \pi, \quad (1.22)$$

kde p je kladný reálný parametr. Předpokládá se, že $F(y)$ je libovolná zobecněná funkce, která je definovaná na kladné ose a pro kterou existuje reálné číslo α takové, že $0 < \alpha < p$ a výraz $\left| \int_{y_1}^{y_2} y^{\alpha-p} F(y) dy \right|$ je stejně omezený pro libovolné $0 < y_1 < y_2$. Obraz $G(z)$ je potom holomorfní na $|\arg z| < \pi$ a existuje $\beta > 0$ tak, že absolutní hodnota $|z^\beta G(z)|$ je omezená v nekonečnu.

Alternativně se na zobrazení (1.22) pohlíží jako na násobnou Laplaceovou transformaci. Jedna z možností, jak definovat inverzní operátor, je proto využití násobného operátoru inverzní Laplaceovy transformace. Jiný způsob nabízí již zmiňovaný článek [3], který dokazuje, že inverzní transformaci k (1.22) lze formulovat pomocí jedno-dimenzionálního integrálu. Předpis pro přechod od G k F uvádějí ve formě

$$F(y) = -\frac{1}{2\pi i} y^p \int_{\mathcal{C}} \frac{G'(yw)dw}{(1+w)^{1-p}}, \quad (1.23)$$

kde \mathcal{C} je křivka v komplexním prostoru, mající začátek i konec v bodě -1 a obíhající počátek ve směru hodinových ručiček.

Srovnáme-li rovnici (1.21), kterou se snažíme vyřešit, a obecný tvar Stieltjesovy transformace, zjistíme, že nás zajímá případ, kdy parametr p je roven 1. V definici (1.23) můžeme snadno provést integraci, zvolíme-li parametr $p = 1$, díky platnosti Cauchyho teorému, který poskytuje správné řešení

$$F(y) = \lim_{\varepsilon \rightarrow 0^+} \frac{1}{2\pi i} (G(-y - i\varepsilon) - G(-y + i\varepsilon)). \quad (1.24)$$

pro $y > 0$.

Funkci $\tilde{u}(t)$ máme určenou na kladné části reálné osy pomocí analytické funkce $u(t)$ (viz (1.21)). Chceme-li na ni použít inverzní operátor (1.24), musíme ji nejprve holomorfně dodefinovat na celém $|\arg t| < \pi$. Holomorfní tvar funkce $\tilde{u}(t)$ můžeme zapsat vzorcem

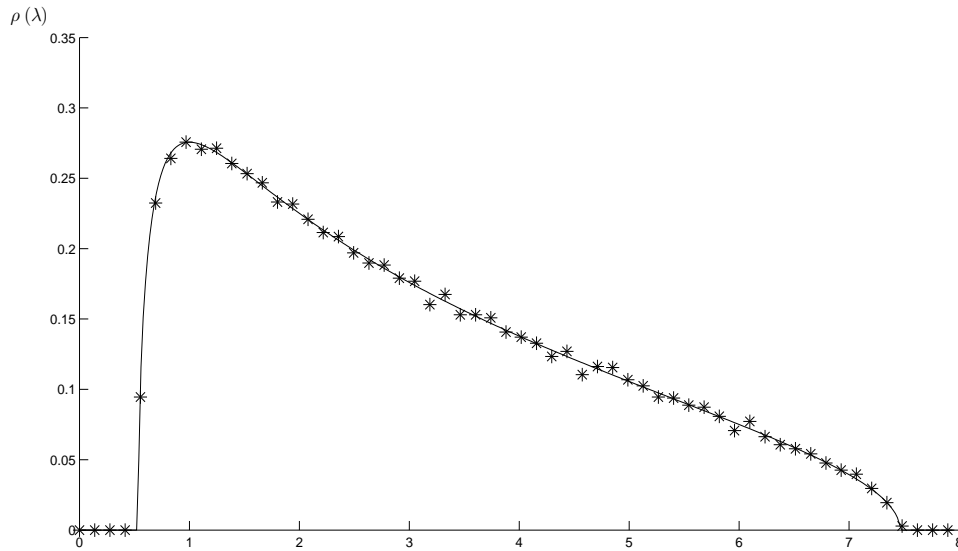
$$\tilde{u}(t) = \frac{-1 - \frac{v(d-1)}{t} + \sqrt{\left[1 + \frac{v(d-1)}{t}\right]^2 + \frac{4v}{t}}}{2v} \quad (1.25)$$

s tím, že hodnota $\tilde{u}(t)$ v bodě 0 je rovna $\frac{1}{v(d-1)}$ je-li $d \neq 1$. Pro $d = 1$ jde limita $\tilde{u}(t)$ zprava v bodě 0 do nekonečna.

Hustotu vlastních čísel $\varrho(\lambda)$ nalezneme jako limitu ve formuli (1.24) při volbě $G(y) \approx \tilde{u}(y)$ a $F(y) \approx \varrho(\lambda)$. Výsledný tvar

$$\varrho(\lambda) = \begin{cases} \frac{1}{2\pi\lambda v} \sqrt{(\lambda_1 - \lambda)(\lambda - \lambda_2)} & \lambda_1 \leq \lambda \leq \lambda_2 \\ 0 & \lambda < \lambda_1 \vee \lambda > \lambda_2 \end{cases}, \quad (1.26)$$

kde $\lambda_{1,2} = v \left(1 \pm \sqrt{d}\right)^2$, je stejný jako v Pastur-Marchenkově předpovědi (1.3), včetně existence Diracovy delta funkce v bodě 0 o velikosti $1 - d$ v případě, že $d < 1$. Důkaz spočívá v dosazení parametrů $d = \frac{1}{c}$ a $v = c\sigma^2$ do vzorce (1.26). Volba parametrů d a v přímo odpovídá předpokladům uvedeným odstavci 1.2.



Obrázek 1.2: Experimentální výsledky (*)¹ proložené grafem funkce $\varrho(\lambda)$ pro $v = 1$ a $d = 3$

Shoda obou výsledků je velmi zajímavá, zvláště, když si uvědomíme, že nalezená hustota není závislá na typu rozdělení elementů v matici x (implicitně jsme předpokládali Gaussovo).

Závěr: Hustota vlastních čísel korelační matice xx^T má tvar (1.26) (viz obrázek 1.2), který nezávisí na počtu zdrojů vln ani na jak vysokých frekvencích vysílají, důležitý je pouze součet šířek jednotlivých vysílacích pásem a tvar výkonového spektra. Rozdělení je nenulové pouze na intervalu určeném λ_1 a λ_2 .

1.6 Speciální případ barevného šumu

Na základě tvaru modifikovaného výkonového spektra libovolného barevného šumu můžeme stanovit rovnice, kterým musí hustota vlastních čísel korelačních matic vyhovovat. Obdržené rovnice se v obecném případě těžko řeší. Proto se je pokusme vyřešit alespoň v nejjednodušších případech.

V reálném experimentu vypadají funkce $v_n(y)$, definované formulí (1.14), jako schodovité funkce (viz odstavce 1.4), protože vyšetřované časové řady jsou konečné. Výkonové spektrum, jenž funkce $v_n(y)$ určuje, je v principu možné zpřesnit dvěma způsoby. Prodloužení pozorovacího okénka je jedna z možností, která není příliš vhodná. Ve skutečnosti při ní spojujeme po sobě jdoucí okénka, která sama o sobě

¹Experimentální výsledky byly získány zpracováním náhodných dat, generovaných počítačem pro $v = 1$ a $d = 3$.

nesou dílčí informaci o studovaném systému za určitou dobu, což může vést k mylným závěrům o stavu probíhajícího procesu.

Mnohem atraktivnější způsob zpřesnění výkonového spektra je zvyšování vzorkovací frekvence měřicího přístroje, jelikož délka pozorovacího okénka může zůstat stejná. Naneštěstí vzorkovací frekvence je základní vlastností měřidla, a proto ji obvykle vůbec nelze měnit, nebo pouze v omezeném rozsahu.

Spokojíme-li se s určitou přesností ve stanovení modifikovaného výkonového spektra, můžeme učinit předpoklad, že se zvyšováním vzorkovací frekvence se spektrum nezmění ve smyslu definice funkcí $v_n(y)$. Posloupnost $\{v_n(y)\}$ je potom konstantní a má za limitu schodovitou funkci.

Korektní odvození hustoty vlastních čísel $\varrho(\lambda)$ probírané v tomto odstavci je po formální stránce velmi složité, nebudeme proto uvádět přesné výpočty u každého kroku, spíše se budeme snažit, aby celkový algoritmus zůstal přehledný.

Dosazením schodovité funkce $v(y)$ do integrálu ve výrazu (1.19) a užitím definice $\tilde{u}(t)$ (viz úprava 1.21) získáme rovnost

$$t\tilde{u}(t) + \sum_{i=1}^K d_i - 1 = \sum_{i=1}^K \frac{d_i}{1 + v_i \tilde{u}(t)} \quad (1.27)$$

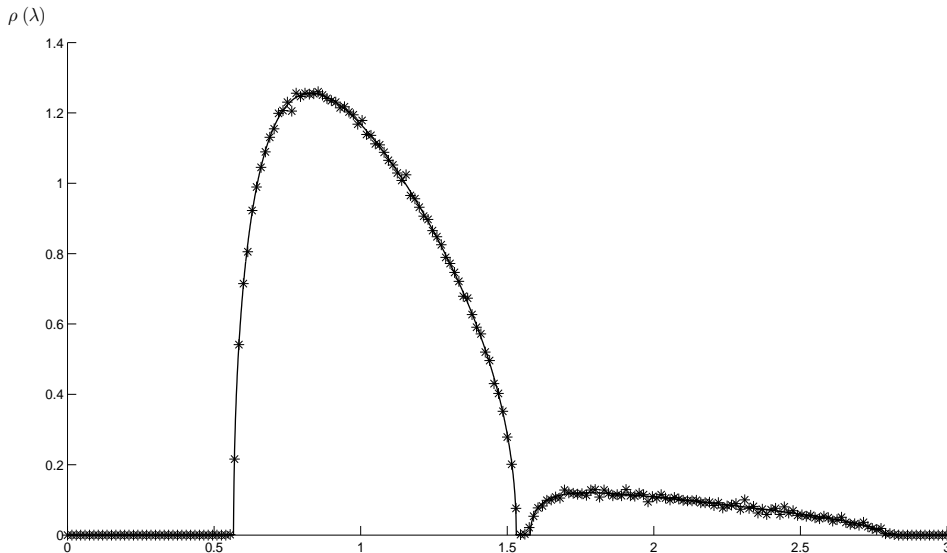
kde K je počet schodů ve funkci $v(y)$, d_i a v_i pak jejich výška a délka.

Řešení rovnice (1.27) vyústí v hledání kořenů polynomu $(K+1)$ -ho stupně v $\tilde{u}(t)$. Výsledný předpis pro $\tilde{u}(t)$, pokud existuje, je v závislosti na parametrech d_i a v_i velmi dlouhý. Z tohoto důvodu ho zde nebudeme uvádět.

Průběh funkce $\tilde{u}(t)$ není závislý na umístění schodů ve $v(y)$, tj. kdybychom zpřeházeli pořadí schodů ve $v(y)$ s tvarem $\tilde{u}(t)$ se nic nestane. Také spojovat schody o stejné výšce do jednoho širšího můžeme beztržně.

Tvrzení Věty 1.4.3 zaručuje, že funkce $u(t)$ bude kladná na pozitivní reálné poloose a v důsledku relace (1.21) také $\tilde{u}(t)$. Pokud budeme chtít použít Stieltjesovu transformaci musíme navíc požadovat holomorfnost na $|\arg t| < \pi$, kterou zařídíme, podobně jako v odstavci 1.5, správným dodefinováním mimo kladnou reálnou poloosu.

Hledanou hustotu vlastních čísel $\varrho(\lambda)$ pro barevný šum, narozdíl od zobecněného bílého šumu, nemůžeme zapsat jediným přehledným vzorcem. Dokonce zvolíme-li hodnoty parametrů



Obrázek 1.3: Experimentální výsledky $(*)^2$ proložené grafem funkce $\varrho(\lambda)$

pro $K = 2$, $v_1 = 1$, $v_2 = \frac{1}{15}$, $d_1 = \frac{1}{10}$, $d_2 = 15$

$$K = 2, v_1 = 1, v_2 = \frac{1}{15}, d_1 = \frac{1}{10}, d_2 = 15, \quad (1.28)$$

nevešel by se přesný vzorec hustoty $\varrho(\lambda)$ na několik řádků. Graf funkce $\varrho(\lambda)$ ve vybraném případě je na obrázku 1.3.

Závěr: Hustota vlastních čísel $\varrho(\lambda)$ korelačních matic xx^T v případě, kdy frekvenční rozklad signálů tvořících matici x odpovídá barevnému šumu, má podobný tvar jako je na obrázku 1.3. Nezávisí na kmitočtu jednotlivých vysílacích pásem (pouze na jejich výkonu úměrnému v_i a šířce d_i), ani na jejich pořadí (viz obdobné úvahy o přehazování schodů ve funkci v). Rozdělení je nenulové pouze na některých intervalech.

²Experimentální výsledky byly získány zpracováním náhodných dat generovaných počítačem pro $K = 2, v_1 = 1, v_2 = \frac{1}{15}, d_1 = \frac{1}{10}, d_2 = 15$.

Kapitola 2

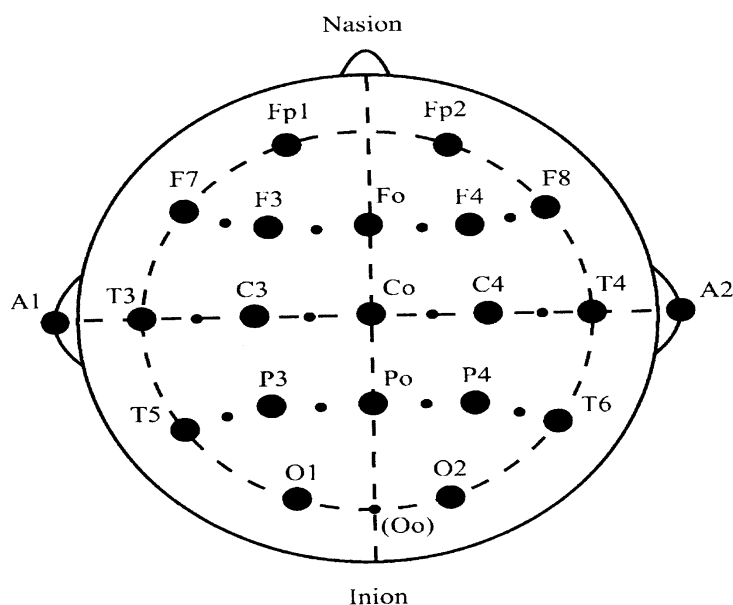
Analýza lidského elektroencefalogramu

Mozek je složený převážně z gliových buněk, které jsou uvnitř, a neuronů, které pokrývají povrch. Obal neuronů je tvořen elektricky nabitou membránou. Elektrický potenciál membrány vzniká při metabolické přeměně látek dodávaných krví (glukóza a kyslík). Neurony jsou propojeny dvěma druhy výběžků, dendrity a neurity neboli axony. Navzájem spolu komunikují prostřednictvím "chemických posílů" neurotransmiterů čili mediatorů.

Biochemicko-elektrický signál je přijímán na membránách dendritů a somatu (těle) buněk. Po neuritu se šíří informace k jiným neuronům. Membránové potenciály jsou jednak pomalé, řádově desítky Hz a rychlé, řádově desítky až stovky Hz. Pomalé nazýváme "stálé" ("standing"), rychle nazýváme "hrotové" ("spiking"). Synchronizací především "standing" potenciálů vzniká elektroencefalogram (EEG).

2.1 Měření EEG

Rozložení elektrod na hlavě není náhodné, řídí se jednoduchým systémem, který navrhl H. Jasper (viz [4]). Příklad jednoho z možných zapojení je na obrázku 2.1.



Obrázek 2.1: Rozmístění elektrod na povrchu lebky s označením jednotlivých elektrod

Spojení mezi dvěma elektrodami je buď bipolární čili diferenční, kdy zesilovací kanál má na vstupu

vždy dvě elektrody, nebo je spojení unipolární čili referenční, kdy k jedné elektrodě (Goldmannově) jsou ostatní připojeny přes určitý odpor. Pro zpracování EEG signálu na počítači se používá unipolární zapojení.

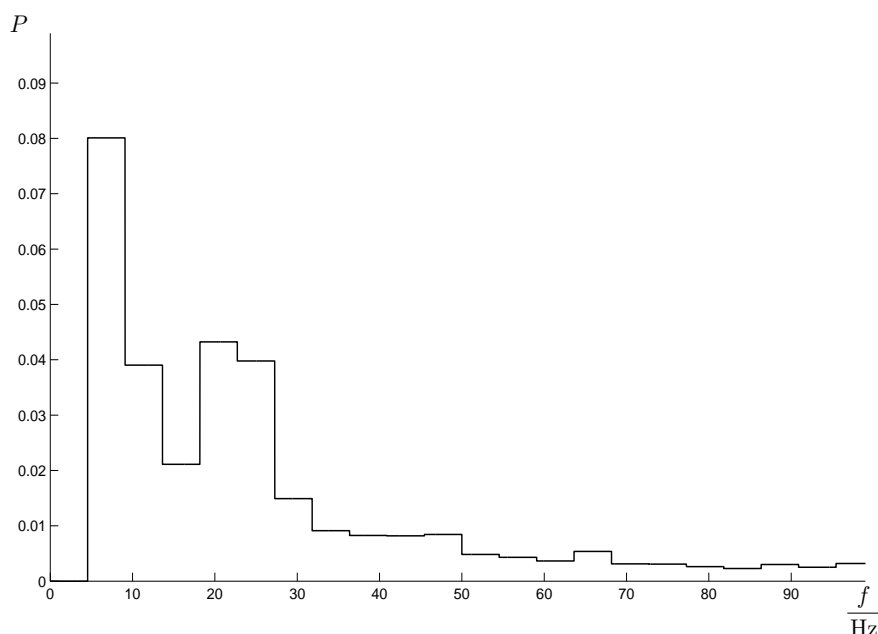
2.2 Frekvenční charakteristika

Změny potenciálu synaptodendritických membrán v nejvyšších vrstvách koncového mozku (případně z jiných částí, ale odtud hlavně) se přenášejí na povrch lebky, kde jsou zaznamenány elektrodami. Při své cestě musí procházet různým prostředím. V závislosti na frekvenci může docházet, buď ke ztrátám energie, nebo k úplnému pohlcení vlny. Nejvyšší frekvence se díky tomuto "organickému filtru" nedostanou na povrch. Proto se dříve myslelo, že v mozku tak rychlé procesy neprobíhají.

Základní aktivitou mozku, který je zdravý, bdělý a zralý (přibližně od 5-7 roku), je tzv. "alfa-a" aktivita, vyznačující se frekvencí v rozsahu 8-13 Hz a amplitudou 30-80 μV . Na druhém místě jmenujme "beta" aktivitu s frekvencí 14-30 Hz a amplitudou 30 až 50 μV . Mezi nejdůležitější patří také "theta" aktivita s frekvencí 4-7,5 Hz a amplitudou do 30 μV . Ostatní druhy aktivity se vyskytují jen zřídka nebo jsou důsledkem nervových poruch.

Výstupy z elektrod putují do zesilovačů, kde kromě zesílení signálu může dojít k odřezání některých frekvencí. Na konec jsou výsledky zaznamenány v počítači ve formě časových řad $\{x_j^i\}_{i=1, j=0}^{i=M, j=n-1}$, kde M je počet elektrod (kanálů) a N určuje délku měření v závislosti na vzorkovací frekvenci f_v (viz vzorec 1.5).

Zpracování dat začneme volbou šířky pozorovacího okénka $n = cM$, kde c je konstanta z Věty 1.4.3. Pomocí diskretní fourierovy transformace zobrazíme časové řady a vytvoříme power spektra (viz odtavec 1.3), která zprůměrujeme přes všechna okénka.



Obrázek 2.2: Výkonové spektrum signálů přijímaných elektrodami při měření EEG

Příklad výkonového spektra získaného z 64-kanálového EEG je na obrázku 2.2. Vidíme, že jednotlivé frekvence nejsou zastoupeny rovnoměrně. Nejvyšší frekvence v rozkladu chybí úplně. Vysvětlení bychom, v souladu s předchozími úvahami, pravděpodobně našli ve vlastnostech prostředí uvnitř lebky a v lebce samotné, nebo v charakteristikách zesilovačů.

2.3 Hustota vlastních čísel korelačních matic

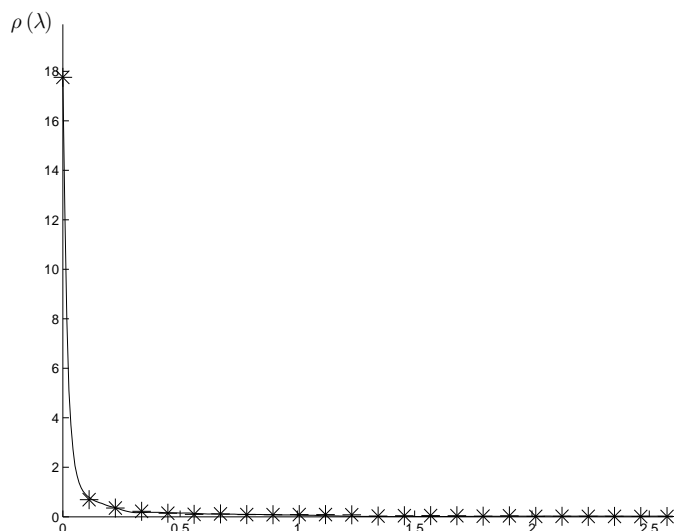
Data z lidského EEG jsme v předchozím odstavci zobrazili diskrétní fourierovou transformací. Obdrželi jsme frekvenční charakteristiku v podobě výkonového spektra.

Spektrální analýza EEG spočívá v konfrontaci zvoleného modelu mozku s naměřenými hodnotami. Náš model je založen na jednoduchém předpokladu, že signál uvnitř lebky je obyčejný bílý šum. Pronikáním na povrch hlavy se z něho stává barevný šum. Stále zůstává náhodným signálem, narozdíl od skutečné elektromagnetické aktivity, která přenáší informaci mezi jednotlivými částmi mozku, respektive částmi celého těla.

Na první pohled se může zdát, že kolísání potenciálu na membránách každého z neuronů nemá globální význam. Shromáždí-li se skupina neuronů, výkazujících podobné změny napětí (vnitřně provázané změny), na jednom místě, stává se z nich samostatná jednotka schopná vykonávat jistou funkci. EEG je tvořeno elektromagnetickou aktivitou právě takových shluků.

Plně náhodný signál, jako je bílý šum, nemůže přenášet informaci. Navržený model mozku proto musí být v rozporu s experimentálně naměřenými daty. Hustota vlastních čísel by mohla být tím ukazatelem, který odhalí nedostatečnost našeho popisu činnosti mozku.

Na základě výsledků odvozených v kapitole 1, jsme schopni nalézt spektrální hustotu korelačních matic barevného šumu, podle průběhu výkonového spektra (viz obrázek 2.2). Společně s experimentálně zjištěným rozdělením vlastních čísel jsou na obrázku 2.3.



Obrázek 2.3: Hustota vlastních čísel korelačních matic z EEG (*) proložená hustotou $\rho(\lambda)$ pro barevný šum s výkonovým spektrem z obrázku 2.2

V oblasti kolem nuly je hustota $\rho(\lambda)$ téměř singulární, přesto je námi zvolená aproximace schopná ji postihnout. Polovina vlastních čísel je velmi malých a druhá polovina je přibližně rovnoměrně rozmístěna na zbytku kladné reálné poloosy.

Závěr: Hustota vlastních čísel korelačních matic, konstruovaných z časových řad pocházejících z elektrod umístěných na hlavě, odpovídá předpovědím učiněným prostřednictvím znalosti výkonového spektra. Na úrovni spektrálních hustot není možné odlišit data z elektroencefalografu od barevného šumu.

Literatura

- [1] W. A. MARCHENKO, L. A. PASTUR: *Distribution eigenvalues of sample covariance matrices*. Math. USSR-Sbornik, 1:457-486, 1967.
- [2] V. L. GIRKO: *Theory of Random Determinants*. Kluwer 1990.
- [3] J. H. SCHWARZ: *The Generalized stieltjes transform and its inverse*. Pasadena, USA, 2004.
- [4] J. FABER: *Elektroencefalografie a psychologie*. ISV, Praha, 2001.