

VAJA 3

Uporaba LSE, PCA in PCR metode za identifikacijo modela

V prilogi so podani izmerjeni podatki za proces oksidacije amoniaka v dušikovo kislino. Meritve so za naslednje spremenljivke: pretok zraka za hlajenje (A_{FLOW}), temperatura vode (T_{H2O}), koncentracija kisline (C_{ACID}) in inverzna vrednost izkoristka (I_{EFF}). Končni cilj naloge je identificirati enačbo s katero je možno izračunati inverzno vrednost izkoristka na podlagi ostalih izmerjenih spremenljivk.

1. VPLIV STANDARDIZACIJE PODATKOV NA REZULTATE:

Preveri in med seboj primerjaj modele (parametre enačbe) dobljene z PCA (implicitni način pridobivanja modela z lastnimi vektorji v obliki $p^*(x-v)=0$) metodo in LSE metodo pri različnih načinih standardizacije podatkov. Upoštevaj, da pri uporabi metode PCA dobiš model v implicitni obliki in jo je potrebno razviti v eksplicitno.

Preizkusi naslednje pristope standardizacije podatkov:

- Standardizacija podatkov na interval [0 1]: $X = (X - \min(X)) / (\max(X) - \min(X))$
- Standardizacija podatkov s standardno deviacijo in odštetim povprečjem:
 $X = (X - \text{mean}(X)) / \text{std}(X)$
- Nestandardizirani podatki.

Za primere a), b) in c) izračunaj model z metodo PCA. Za primer a) in b) je potrebno dobljene parametre transformirati nazaj v originalni prostor.

Dobljene rezultate predstavi v poročilu in jih komentiraj. Ali ima standardizacija vpliv na parametre? Preveri kakšna je porazdelitev napake. Kako se spreminjajo parametri glede na uporabljeno metodo. Za LSE metodo izračunajte varianco parametrov (teoretično).

2. PROBLEM KOLINEARNOSTI:

Podanim meritvam dodajte še eno meritev, ki je odvisna od temperature:

$$x = 2 * T_{H2O} + 6 + 0.1 * \text{randn}(\text{length}(T_{H2O}), 1).$$

Izračunaj parametre po PCR metodi. Pri PCR metodi se podatki prek PCA metode preslikajo v drug prostor, naredi se LSE in preslika parametre nazaj v originalni prostor. Identifikacijo zaženite večkrat in opazujte kaj se dogaja s parametri. Postopek ponovite z uporabo LSE metode.

Kaj opaziš pri dobljenih parametrih? Kako bi lahko ugotovil kateri spremenljivki sta kolinearni? Kako se spreminjajo parametri glede na uporabljeno metodo (PCR in LSE).

Za LSE metodo izračunajte varianco parametrov (teoretično). Ocenite varianco parametrov za LSE in PCR iz poskusov.

Predmet: Inteligentni sistemi za podporo odločanju

V poročilu opišite metodologijo, ki ste jo uporabili in predstavite dobljene rezultate.