

PCA - Inspiration til matematikfagkobling

Udvidelse af forsøget "Analyse af benzoxazinoide i brød"

Af Rasmus Knokgaard og Nicole Koefoed fra Hvidovre Gymnasium og HF
og Stine Krogh Steffensen fra Institut for Agroøkologi, AU, Flakkebjerg

Eleven har forberedt før øvelsen:

1. Elev og lærer har udvalgt hvilke(n) del(e) af den udvidede efterbehandling eleven skal arbejde med.
2. Eleven har forberedt sig grundigt på ved at læse den valgte litteratur.
3. Eleven har installeret LatentiX eller et andet program til PCA-analyse, og sat sig ind i hvordan man bruger det.

Lidt om øvelsen og dens formål

Som forsker kan man ikke nøjes med bare at være rigtigt god til ét fag. Man skal kunne bruge værktøjer fra mange forskellige fag. I forsøget "Analyse af benzoxazinoide i brød" er det i høj grad kemiens værktøjer man bruger. Men når de indsamlede data skal behandles, skal man have fat i matematikkens værktøjer. Et sådant værktøj er PCA. PCA er et multivariat statistisk analyseværktøj der kan hjælpe en med at finde interessante tendenser i sine data. Med det kan man se på sine data på en mere visuel og overskuelig måde.

Formålet med denne udvidelse til "Analyse af benzoxazinoide i brød" at gå endnu dybere ned efterbehandlingen af forskningsdata end man normalt gør i en almindelig kemirapport. Dette er ikke en færdisyet matematikopgave, hvor man bare kan begynde med første opgave og slutte med sidste. Dette skal ses som en kilde til inspiration til hvad man blandt andet kan vælge at arbejde med hvis man vil arbejde med fagkombinationen kemi/matematik. Denne del af øvelsen står elev og lærer altså selv for. Men PCA-analysen kan påbegyndes under opholdet på Institut for Agroøkologi, Flakkebjerg. Ekstra brøddata kan udleveres for at få nok data til en PCA-analyse.

Inspiration til den udvidede efterbehandling:

Apparatur: Computer med tekstbehandling, regneark og program til at lave PCA-analyse (f.eks. LatentiX).

Arbejdsformer:

Man kan vælge en af to tilgange til materialet eller vægte deres bidrag i opgaven efter interesse og tidligere undervisningsforløb. Den ene er teoretisk og den anden giver mulighed for en mere praktisk tilgang til matematikfaget.

1. Lineær algebra.

Alt efter interesse, og eventuelle tidligere undervisningsforløb, kan man fordybe sig i den teoretiske side af PCA. Det er for meget at kræve at eleven har en fuld forståelse af den matematiske baggrund, men man kunne arbejde med nogle af de begreber inden for lineær algebra som PCA-analysen bygger på, f.eks.:

- Linear kombination
- En basis for et vektorrum (row space and column space)
- Underrum (et span)
- Basisskift (ikke for et generelt vektorrum)
- Ortogonalitet
- Correlation matrix
- Covariancematrix
- Ortogonale underrum
- Egenverdier og egenvektorer

2. PCA

Her kan man lægge vægten på flere forskellige aspekter:

- **Formidling.** En beskrivelse af, hvad der foregår, når man udfører en PCA, uden decideret at gå i detaljer med selve matematikken. (Her kunne man inddrage danskfaget også, hvis man er interesseret i en formidlingsopgave.)
- **Diskussion.** Er det problemfrit at manipulere data?
- **Statistik.** Her er talrige muligheder, som må afhænge af interesse og eventuelle tidligere undervisningsforløb. Det kan være vilkårligt teoretisk, men også mere diskuterende, i forhold til hvilke muligheder man har, for at trække information ud af sine data.

Litteratur og materiale:

Steven J. Leon (2006). Linear Algebra with applications (seventh edition)

Der findes mange glimrende informationer online. Et lille udpluk:

Online videoer:

http://www.youtube.com/watch?v=UUxIXU_Ob6E og <http://www.youtube.com/watch?v=26YhtSji1qc>

<http://www.youtube.com/watch?v=f9mZ8dSrVJA>

Tekster:

<http://www.kemi-online.dk/?p=27082>

<http://www.brainmapping.org/NITP/PNA/Readings/pca.pdf>

http://www.ce.yildiz.edu.tr/personal/songul/file/1097/principal_components.pdf

people.maths.ox.ac.uk/richardsonm/SignalProcPCA.pdf

Til at udføre PCA findes også adskillige gratis programmer online - f.eks LatentiX, XLStat eller Multibase