Algebraic Eigenvalue Problem, J.H Wilkinson

Να υπολογιστεί ένας στοιχειώδης Ερμιτιανός πίνακας σε αριθμητική κινητής υποδιαστολής

Ένας πίνακας P καλείται στοιχειώδης Ερμιτιανός αν είναι unitary της μορφής

$$P = I - 2ww^\intercal = \left[\begin{array}{c|c} I & 0 \\ \hline 0 & I - 2vv^\intercal \end{array}\right]$$

όπου

$$v^\intercal v = w^\intercal w = 1.$$

Έστω $x \in \mathbb{R}^n$, να βρεθεί πίνακας P τέτοιος ώστε πολλαπλασιαζόμενος από αριστερά με το x να μηδενίζει τα στοιχεία του $r+2,\ r+3,\dots,n$ του x αφήνοντας τα r πρώτα στοιχεία αμετάβλητα.

Χωρίς βλάβη της γενικότητας μπορούμε να υποθέσουμε ότι $r = 0$.

Υπολογισμός

Αναζητούμε πίνακα P τέτοιον ώστε

$$Px = ke_1.$$

Έστω

$$S^2 = x_1^2 + \ldots + x_n^2\,,$$

αφού P unitary έχουμε ότι

$$k = \pm S\,.$$

Η αρχική εξίσωση μας δίνει

$$\begin{cases} x_1 - 2w_1(w^\intercal x) = \pm S, & (i = 1) \\ x_i - 2w_i(w^\intercal x) = 0, & i = 2,\dots,n \end{cases}$$

επομένως

$$2Kw_1 = x_1 \mp S, \quad 2Kw_i = x_i,\ i = 2,\dots,n\,,$$

όπου

$$K = w^\intercal x\,.$$

Απο την τελευταία σχέση (προσθέτοντας και διαιρώντας δια 2) έχουμε ότι

$$2K^2 = S^2 \mp x_1 S\,.$$

Οπότε αν θέσουμε

$$u^\intercal = (x_1 \mp S, x_2, \dots, x_n),$$

τότε

$$w = u/(2K),$$

και

$$P = I - ww^\intercal = uu^\intercal/(2K^2)\,.$$

Αριθμητικό παράδειγμα (single precision)

Έστω $n = 3$ και

In [1]:

```
using LinearAlgebra

@show x = [0.5123f0; 0.6147f-3; 0.5135f-3];
@show S2 = sum(x.^2);
@show S = sqrt(S2);
```

```
x = [0.5123f0; 0.0006147f0; 0.0005135f0] = Float32[0.5123, 0.0006147, 0.00
05135]
S2 = sum(x .^ 2) = 0.26245195f0
S = sqrt(S2) = 0.51230067f0
```

Έστω ότι επιλέγουμε $k = -S$ (stable choice)

In [2]:

```
# K2 = x[1]^2 + x[2]^2 + x[3]^2 + x[1]*S
@show K2 = S2 + x[1]*S;
@show u1 = x[1]+S;
@show u = [u1; x[2]; x[3]];
print("signs : ", sign(x[1]), ", ", sign(S));
```

```
K2 = S2 + x[1] * S = 0.5249036f0
u1 = x[1] + S = 1.0246007f0
u = [u1; x[2]; x[3]] = Float32[1.0246007, 0.0006147, 0.0005135]
signs : 1.0, 1.0
```

In [4]:

```
P = I(3)-u*u'/K2;
display(P)
display(P*P')
@show opnorm(P*P'-I(3),Inf);
```

```
3×3 Array{Float32,2}:
 -0.999999    -0.00119988  -0.00100234
 -0.00119988   0.999999    -6.01346f-7
 -0.00100234  -6.01346f-7   1.0

3×3 Array{Float32,2}:
 1.0          2.53507f-10  2.32831f-10
 2.53507f-10  1.0          5.68434f-14
 2.32831f-10  5.68434f-14  1.0

opnorm(P * P' - I(3), Inf) = 4.773235f-7
```

Έστω ότι επιλέγουμε $k = +S$ (unstable choice)

In [5]:

```
@show K2 = S2 - x[1]*S;
@show u1 = x[1]+(-S);
@show u = [u1; x[2]; x[3]];
println("values : ", x[1], ", ", (-S))
println("signs : ", sign(x[1]), ", ", sign(-S));
```

```
K2 = S2 - x[1] * S = 2.9802322f-7
u1 = x[1] + -S = -6.556511f-7
u = [u1; x[2]; x[3]] = Float32[-6.556511f-7, 0.0006147, 0.0005135]
values : 0.5123, -0.51230067
signs : 1.0, -1.0
```

In [7]:

```
P = I(3)-u*u'/K2;
display(P)
display(P*P')
@show opnorm(P*P'-I(3),Inf);
```

```
3×3 Array{Float32,2}:
 0.999999     0.00135234   0.0011297
 0.00135234  -0.267875    -1.05914
 0.0011297   -1.05914      0.115229

3×3 Array{Float32,2}:
  1.0          -0.000206431  -0.000172445
 -0.000206431   1.19354       0.161675
 -0.000172445   0.161675      1.13506

opnorm(P * P' - I(3), Inf) = 0.3554183f0
```

Αριθμητικό παράδειγμα σε αριθμητική με mantissa 13 bits (περίπου 4 ψηφία στο δεκαδικό)

In [8]:

```
using LinearAlgebra
setprecision(13)

@show x = [BigFloat("0.5123e0"); BigFloat("0.6147e-3"); BigFloat("0.5135e-3")];
@show S2 = sum(x.^2);
@show S = sqrt(S2);
```

```
x = [BigFloat("0.5123e0"); BigFloat("0.6147e-3"); BigFloat("0.5135e-3")] =
BigFloat[0.51233, 0.00061464, 0.00051355]
S2 = sum(x .^ 2) = 0.26245
S = sqrt(S2) = 0.51233
```

Έστω ότι επιλέγουμε $k = -S$ (stable choice)

In [9]:

```
# K2 = x[1]^2 + x[2]^2 + x[3]^2 + x[1]*S
@show K2 = S2 + x[1]*S;
@show u1 = x[1]+S;
@show u = [u1; x[2]; x[3]];
```

```
K2 = S2 + x[1] * S = 0.5249
u1 = x[1] + S = 1.0247
u = [u1; x[2]; x[3]] = BigFloat[1.0247, 0.00061464, 0.00051355]
```

In [11]:

```
P = I(3)-u*u'/K2;
display(P)
display(P*P')
@show opnorm(P*P'-I(3),Inf);
```

```
3×3 Array{BigFloat,2}:
 -1.0          -0.00119972   -0.00100255
 -0.00119972    1.0          -6.01402e-07
 -0.00100255   -6.01402e-07   1.0

3×3 Array{BigFloat,2}:
 1.0          6.02881e-10  0.0
 6.02881e-10  1.0          0.0
 0.0          0.0          1.0

opnorm(P * P' - I(3), Inf) = 6.0288e-10
```

Έστω ότι επιλέγουμε $k = +S$ (unstable choice)

In [12]:

```
@show K2 = S2 - x[1]*S;
@show u1 = x[1]-S;
@show u = [u1; x[2]; x[3]];
```

```
K2 = S2 - x[1] * S = 0.0
u1 = x[1] - S = 0.0
u = [u1; x[2]; x[3]] = BigFloat[0.0, 0.00061464, 0.00051355]
```

In [14]:

```
P = I(3)-u*u'/K2;
display(P)
display(P*P')
@show opnorm(P*P'-I(3),Inf);
```

```
3×3 Array{BigFloat,2}:
 NaN  NaN   NaN
 NaN  -Inf  -Inf
 NaN  -Inf  -Inf
```

```
3×3 Array{BigFloat,2}:
 NaN  NaN  NaN
 NaN  NaN  NaN
 NaN  NaN  NaN
```

```
opnorm(P * P' - I(3), Inf) = NaN
```