Article

Network pharmacology of medicinal attributes and functions of Chinese herbal medicines: (III) Canonical correlation functions between attribute classes and linear eigenmodels of Chinese herbal medicines

WenJun Zhang
School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, China; International Academy of Ecology and Environmental Sciences, Hong Kong
E-mail: zhwj@mail.sysu.edu.cn, wjzhang@iaees.org

Received 10 April 2017; Accepted 28 April 2017; Published 1 September 2017

Abstract
In present study I used the data from CHM-DATA, the interactive database of 1127 Chinese herbal medicines. Canonical correlation functions were determined for taste attribute class (7 taste attributes), medicinal property class (5 medicinal properties), chemical composition class (22 chemical composition categories), meridians and collaterals class (12 meridians and collaterals), and medicinal function class (77 medicinal functions). Linear eigenmodels were also developed for Chinese herbal medicines. Theoretically the attribute values of any Chinese herbal medicines meet the corresponding linear eigenmodel. Matlab codes for canonical correlation analysis and linear eigenmodel were given. Finally, the canonical correlation network for attribute classes of Chinese herbal medicines was constructed.

Keywords Chinese herbal medicine; medicinal function; attribute; canonical correlation; eigenmodel.

1 Introduction
The single drug-single target-single disease view in traditional western medicine (Hopkins, 2007, 2008; Budovsky and Fraifeld, 2012) met various problems over the past 20 years (Zhang, 2016, 2017a-b). Traditional Chinese Medicine takes biological network regulation as the theoretical basis, and thus provides a new thinking and new approach for drug design and disease treatment. However, the theory of Traditional Chinese Medicine has developed so slowly in the past thousands of years. So far we still lack of fundamental research on Chinese herbal medicines, which greatly retards the development and practice of the theory of Chinese herbal medicines. For this reason, Zhang (2017a) collected a total of 1127 Chinese herbal medicines mainly with recorded chemical composition, and calculated the basic statistics of medicinal attributes and functions, e.g., totals, frequencies or probabilities, percentages, etc., on the basis of total population of medicines and families. Thereafter, four relational networks, i.e., the networks for medicinal attributes and
functions, for chemical composition and meridians and collaterals, for meridians and collaterals and medicinal functions, and for meridians and collaterals were constructed based on the significant point correlations (Zhang, 2017b). Network analysis indicated that the former three ones are scale-free complex networks and node degrees of the four networks followed power-law distribution. Detailed between-attribute relationships and medicinal mechanisms were revealed (Zhang, 2017b).

Based on the previous studies (Zhang, 2017a, b), this study will further determine the canonical correlations between attribute classes and develop the standard models of Chinese herbal medicines, in order to lay a foundation for further studies.

2 Material and Methods

2.1 Methods

2.1.1 Canonical Correlation Analysis

In present study, canonical correlation analysis is used to determine the correlation between two attribute classes (Zhang and Fang, 1982; Qi and Xu, 2009). Suppose the attribute class \( x \) has \( m \) attributes \( x_1, x_2, \ldots, x_m \), and the attribute class \( y \) has \( p \) attributes \( y_1, y_2, \ldots, y_p \), \( m \leq p \). We want to analyze the correlation by determining the correlation between \( ux^T \) and \( vy^T \), where \( u = (u_1, u_2, \ldots, u_m) \), \( v = (v_1, v_2, \ldots, v_p) \). The degree of correlation between \( ux^T \) and \( vy^T \) changes with different \( u \) and \( v \). We need to determine \( u \) and \( v \), such that the linear correlation between \( ux^T \) and \( vy^T \) is the strongest. First, assume there are \( n \) medicines, and the raw data are as follows

\[
\begin{align*}
  x &= (x_{ij}), & j = 1, 2, \ldots, m \\
  y &= (y_{ij}), & j = 1, 2, \ldots, p \\
  i &= 1, 2, \ldots, n
\end{align*}
\]

In present study, \( x_{ij} \) and \( y_{ij} \) take 0 or 1. Let \( x_{ij}^* = x_{ij} - x_{bar,j} \), \( y_{ij}^* = y_{ij} - y_{bar,j} \), where

\[
\begin{align*}
  x_{bar,j} &= \frac{\sum_{i=1}^{n} x_{ij}}{n}, & j = 1, 2, \ldots, m \\
  y_{bar,j} &= \frac{\sum_{i=1}^{n} y_{ij}}{n}, & j = 1, 2, \ldots, p
\end{align*}
\]

Calculate

\[
\begin{align*}
  e_{ij} &= \sum_{k=1}^{n} x_{ik} x_{kj}/n, & i, j = 1, 2, \ldots, m \\
  f_{ij} &= \sum_{k=1}^{n} y_{ik} y_{kj}/n, & i, j = 1, 2, \ldots, p \\
  g_{ij} &= \sum_{k=1}^{n} x_{ik} y_{kj}/n, & i = 1, 2, \ldots, m; j = 1, 2, \ldots, p \\
  h_{ij} &= \sum_{k=1}^{n} y_{ik} x_{kj}/n, & i = 1, 2, \ldots, p; j = 1, 2, \ldots, m
\end{align*}
\]

Let \( E = (e_{ij}) \), \( F = (f_{ij}) \), \( G = (g_{ij}) \), \( H = (h_{ij}) \). Determine eigenvalues \( l_1^2, l_2^2, \ldots, l_m^2 \), and the corresponding eigenvector pairs \( u_1, v_1; u_2, v_2; \ldots; u_m, v_m \),

\[
\begin{align*}
  (E^* G F^* H) u &= 0 \\
  (F^* H G^* F) v &= 0
\end{align*}
\]

where \( I \) is the unit matrix. Finally, the canonical correlation coefficients (absolute values) are \( l_1, l_2, \ldots, l_m \), and the canonical attribute pairs, or correlation functions are obtained as
2.1.2 Linear eigenmodels

For the above \((u_i, v_i), i=1, 2, ..., m\), develop linear regression with \(u_i\) (or \(v_i\)) as independent variable, and \(v_i\) (or \(u_i\)) as dependent variable, and choose these models with statistic significance. For example

\[ u_i = a + b \cdot v_i \]

Consequently, we achieve a linear eigenmodel of Chinese herbal medicines as the following

\[ \sum_{k=1}^{m} u_{ik} x_k = a + b \sum_{k=1}^{p} v_{ik} y_k \]

In a statistic sense, the attribute values of any Chinese herbal medicine meet the corresponding linear eigenmodels.

The following are the Matlab codes for canonical correlation analysis and linear eigenmodel calculation, CanonicalCorreAnaly.m

```matlab
m=input('Input the number of variables x: ');
p=input('Input the number of variables y: ');
if (m>p) disp('Variables x should be less than variables y'); pause; end
file=input('Input the excel file name of data, e.g., cano.xls. The first m columns are for variables x and the followed p columns are for variables y: ','s');
xy=xlsread(file);
n=size(xy,1);
x=xy(:,1:m);
y=xy(:,m+1:m+p);
xb=zeros(1,m); yb=zeros(1,p); sigx=zeros(m); sigy=zeros(p);
sigxy=zeros(m,p); sigyx=zeros(p,m); mat1=zeros(m); mat2=zeros(p);
u=zeros(m); v=zeros(p); val1=zeros(m); val2=zeros(p);
xb=mean(x);
yb=mean(y);
for i=1:m
    x(:,i)=x(:,i)-xb(i);
end;
for i=1:p
    y(:,i)=y(:,i)-yb(i);
end;
for i=1:m
    for j=1:m
        sigx(i,j)=0;
    end;
end;
```

IAEES www.iaeex.org
for k=1:n
    sigx(i,j)=sigx(i,j)+x(k,i)*x(k,j);
end
    sigx(i,j)=sigx(i,j)/n;
end; end
for i=1:p
for j=1:p
    sigy(i,j)=0;
for k=1:n
    sigy(i,j)=sigy(i,j)+y(k,i)*y(k,j);
end
    sigy(i,j)=sigy(i,j)/n;
end; end
for i=1:m
for j=1:p
    sigxy(i,j)=0;
for k=1:n
    sigxy(i,j)=sigxy(i,j)+x(k,i)*y(k,j);
end
    sigxy(i,j)=sigxy(i,j)/n;
end; end
    sigx=sigx^(-1);
    sigy=sigy^(-1);
    mat1=sigx*sigxy*sigy*sigyx;
    mat2=sigy*sigyx*sigx*sigxy;
[u,val1]=eig(mat1);
[v,val2]=eig(mat2);
for i=1:m
    p2(i)=i;
end
for i=1:m-1
    k=i;
for j=i:m-1
    if (val1(j+1,j+1)>val1(k,k)) k=j+1; end
end
    i2=p2(i); p2(i)=p2(k); p2(k)=i2;
    l=val1(i,i); val1(i,i)=val1(k,k); val1(k,k)=l;
end
iss="u';
for k=1:m
iss=strcat(iss,'No. ',num2str(k),' canonical correlation coefficient: ',num2str(round(sqrt(val1(k,k))*10000)/10000.00),'
');
for i=1:m
e1=num2str(i);
if (u(i,p2(k))>0) e2=num2str(round(u(i,p2(k))*10000)/100000.00);
elseif (u(i,p2(k))<0) e2=num2str(round(abs(u(i,p2(k)))*10000)/100000.00);
end
if (u(i,p2(k))>0) iss=strcat(iss,'+',e2,'x',e1);
elseif (u(i,p2(k))<0) iss=strcat(iss,'-',e2,'x',e1);
end
iss=strcat(iss,'
');
iss=strcat(iss,'v',num2str(k),'=');
for i=1:p
e1=num2str(i);
if (v(i,p2(k))>0) e2=num2str(round(v(i,p2(k))*10000)/100000.00);
elseif (v(i,p2(k))<0) e2=num2str(round(abs(v(i,p2(k)))*10000)/100000.00);
end
if (v(i,p2(k))>0) iss=strcat(iss,'+',e2,'y',e1);
elseif (v(i,p2(k))<0) iss=strcat(iss,'-',e2,'y',e1);
end
iss=strcat(iss,'
Linear regression between u',num2str(k),' and', ' v',num2str(k),'
');
for j=1:n
uxx(j)=x(j,:)*u(:,k);
vyy(j)=y(j,:)*v(:,k);
end
for j=1:2
if (j==1) xx=uxx'; yy=vyy';
else xx=vyy';yy=uxx';
end
[bb,bint,rr,rint,stats]=regress(yy,[ones(n,1) xx]);
if (j==1) iss=strcat(iss,'u',num2str(k),'=',num2str(bb(1)),'+',num2str(bb(2)),'*','v',num2str(k),'
');
else iss=strcat(iss,'v',num2str(k),'=',num2str(bb(1)),'+',num2str(bb(2)),'*','u',num2str(k),'
');
iss=strcat(iss,'Pearson r=',num2str(sign(bb(2))*sqrt(stats(1))),' p=',num2str(stats(3)),'
');
end
end
end
iss=strcat(iss,'u');
end
fprintf(iss)
2.2 Data source

I used the interactive database of eight tables, CHM-DATA Version 1.0 (Zhang, 2017a, b), with 1127 Chinese herbal medicines mainly having recorded chemical composition, of which 210 families and approximately 2000 species of medicinal plants and fungi were involved, which account for approximately 1/5 of medicinal plants and fungi in China. Among them, medicinal plants accounted for 98.94%, and medicinal fungi accounted for 1.06%. The list included the most commonly used or important Chinese herbal medicines. The data with missing attribute values are ignored in the study. Finally, the taste attribute class (7 taste attributes, Table 1), medicinal property class (5 medicinal properties, Table 2), chemical composition class (22 chemical composition categories, Table 3), meridians and collaterals class (12 meridians and collaterals (Gui Jing), Table 4), and medicinal function class (77 medicinal functions (Gong Xiao), Table 5), were used for further analysis.

<table>
<thead>
<tr>
<th>Table 1 Taste attribute class.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taste</td>
</tr>
<tr>
<td>味</td>
</tr>
<tr>
<td>x</td>
</tr>
<tr>
<td>y</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2 Medicinal property class.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Property</td>
</tr>
<tr>
<td>性</td>
</tr>
<tr>
<td>x</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3 Chemical composition class.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemical composition Categories</td>
</tr>
<tr>
<td>成份</td>
</tr>
<tr>
<td>Y</td>
</tr>
<tr>
<td>Chemical composition Categories</td>
</tr>
<tr>
<td>成份</td>
</tr>
<tr>
<td>Y</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4 Meridians and collaterals class.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meridians &amp; Collaterals</td>
</tr>
<tr>
<td>归经</td>
</tr>
<tr>
<td>x</td>
</tr>
<tr>
<td>Meridians &amp; Collaterals</td>
</tr>
<tr>
<td>归经</td>
</tr>
<tr>
<td>x</td>
</tr>
<tr>
<td>Function</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>Clean liver, relax liver, bright eyes or eliminate eye screens</td>
</tr>
<tr>
<td>功效</td>
</tr>
<tr>
<td>y</td>
</tr>
<tr>
<td>Activate water metabolism or excrete water</td>
</tr>
<tr>
<td>功效</td>
</tr>
<tr>
<td>y</td>
</tr>
<tr>
<td>Anti-asthma</td>
</tr>
<tr>
<td>功效</td>
</tr>
<tr>
<td>y</td>
</tr>
<tr>
<td>Improve digestion</td>
</tr>
<tr>
<td>功效</td>
</tr>
<tr>
<td>y</td>
</tr>
<tr>
<td>Tonify blood</td>
</tr>
<tr>
<td>功效</td>
</tr>
<tr>
<td>y</td>
</tr>
<tr>
<td>Antihypertension</td>
</tr>
<tr>
<td>功效</td>
</tr>
<tr>
<td>y</td>
</tr>
<tr>
<td>Inhibit or break energy flow (Qi)</td>
</tr>
<tr>
<td>功效</td>
</tr>
<tr>
<td>y</td>
</tr>
<tr>
<td>Clear away heat</td>
</tr>
<tr>
<td>功效</td>
</tr>
<tr>
<td>y</td>
</tr>
<tr>
<td>Dispel endogenous cold</td>
</tr>
<tr>
<td>功效</td>
</tr>
</tbody>
</table>
3 Results and Analysis

3.1 Meridians and collaterals class vs. chemical composition class

The calculated canonical correlations and linear eigenvectors for meridians and collaterals class (Table 4) vs. chemical composition class (Table 3) are as follows (only the first three groups of results are given)

No.1 canonical correlation coefficient: 0.3621

\[ u_1 = 0.1109x_1 + 0.1073x_2 + 0.06532x_3 + 0.0079x_4 + 0.16885x_5 + 0.12342x_6 + 0.06532x_7 - 0.00023x_8 - 0.16885x_9 - 0.14579x_{10} + 0.14219x_{11} + 0.92625x_{12} \]

\[ v_1 = -0.13876y_1 + 0.02522y_2 + 0.16501y_3 + 0.05993y_4 + 0.28537y_5 + 0.14131y_6 + 0.13212y_7 + 0.10196y_{10} + 0.153y_{12} + 0.15954y_{10} + 0.03951y_{12} + 0.13407y_{10} + 0.15844y_{12} + 0.00407y_{12} + 0.01791y_{14} + 0.18414y_{12} + 0.28158y_{14} + 0.08924y_{12} + 0.12594y_{14} + 0.12925y_{12} + 0.05717y_{14} + 0.06036y_{12} + 0.13855y_{12} \]

Linear regression between \( u_1 \) and \( v_1 \)

\[ u_1 = -2.2731 \times 10^{-17} + 0.81879v_1 \]

Linear regression between \( v_1 \) and \( u_1 \)

\[ v_1 = 6.7937 \times 10^{-19} + 0.16017u_1 \]

Pearson \( r = 0.36214 \), \( p = 0 \)

No.2 canonical correlation coefficient: 0.272

\[ u_2 = 0.03819x_1 - 0.03038x_2 + 0.05457x_3 + 0.22421x_4 + 0.03326x_5 + 0.08942x_6 + 0.08308x_7 + 0.00758x_8 - 0.02874x_9 + 0.4307x_{10} + 0.02628x_{12} \]

\[ v_2 = -0.065y_1 - 0.03642y_2 - 0.11035y_3 - 0.44763y_4 + 0.20343y_5 + 0.15458y_6 + 0.09163y_7 - 0.33568y_8 - 0.04326y_9 + 0.14556y_{10} + 0.12594y_{12} + 0.05717y_{14} + 0.0476y_{12} + 0.01791y_{14} + 0.28158y_{12} - 0.08924y_{14} - 0.09163y_{12} + 0.66085y_{14} - 0.06079y_{12} + 0.02648y_{14} + 0.20732y_{12} + 0.0557y_{14} + 0.26351y_{12} + 0.165y_{14} \]

Linear regression between \( u_2 \) and \( v_2 \)

\[ u_2 = 1.7621 \times 10^{-17} + 0.40353v_2 \]

Linear regression between \( v_2 \) and \( u_2 \)

\[ v_2 = -4.848 \times 10^{-10} + 0.18337u_2 \]

Pearson \( r = 0.27202 \), \( p = 1.4544 \times 10^{-14} \)

No.3 canonical correlation coefficient: -0.2564

\[ u_3 = 0.20417x_1 + 0.46968x_2 + 0.03453x_3 + 0.01421x_4 + 0.18606x_5 - 0.1553x_6 + 0.24551x_7 + 0.05308x_8 + 0.23747x_9 - 0.23558x_{10} + 0.59032x_{12} + 0.39259x_{12} \]

\[ v_3 = 0.03547y_1 + 0.31571y_2 + 0.08208y_3 - 0.33568y_4 + 0.14556y_5 + 0.25362y_6 + 0.6057y_7 + 0.09551y_8 + 0.27246y_{10} + 0.04342y_{12} + 0.25632y_{10} + 0.1215y_{12} + 0.18414y_{14} + 0.06771y_{12} + 0.0101y_{14} + 0.09705y_{12} + 0.11578y_{14} + 0.20692y_{12} + 0.15564y_{14} + 0.15817y_{12} + 0.07883y_{14} \]

Linear regression between \( u_3 \) and \( v_3 \)
\[ u = 3.5186 \times 10^{-17} - 0.31822 \cdot v \]

Linear regression between \( v \) and \( u \)
\[ v = 2.3784 \times 10^{-17} - 0.2066 \cdot u \]
Pearson \( r = -0.25641, p = 4.6907 \times 10^{-13} \).

All canonical correlations and linear eigenmodels are statistically significant with \( p \leq 4.6907 \times 10^{-13} \).

### 3.2 Medicinal property class vs. taste attribute class

The calculated canonical correlations and linear eigenmodels for medicinal property class (Table 2) vs. taste attribute class (Table 1) are as follows (only the first three groups of results are given)

No.1 canonical correlation coefficient: 0.518
\[ u = -0.17722x_1 - 0.07275x_2 + 0.60466x_4 - 0.76488x_5 \]
\[ v = 0.4475y_1 + 0.80495y_2 - 0.03154y_3 - 0.33218y_4 + 0.0044y_5 + 0.071y_6 - 0.18812y_7 \]
Linear regression between \( u \) and \( v \)
\[ u = 1.8564 \times 10^{-17} + 0.76742 \cdot v \]
Linear regression between \( v \) and \( u \)
\[ v = 1.3387 \times 10^{-17} + 0.34969 \cdot u \]
Pearson \( r = 0.51804, p = 0 \)

No.2 canonical correlation coefficient: 0.2801
\[ u = -0.5943x_1 - 0.18773x_2 - 0.0663x_3 - 0.45966x_4 - 0.62919x_5 \]
\[ v = 0.01513y_1 + 0.16663y_2 + 0.38585y_3 + 0.61173y_4 + 0.34056y_5 + 0.55242y_6 - 0.16663y_7 \]
Linear regression between \( u \) and \( v \)
\[ u = 1.5874 \times 10^{-17} + 0.36872 \cdot v \]
Linear regression between \( v \) and \( u \)
\[ v = 2.5131 \times 10^{-17} + 0.2128 \cdot u \]
Pearson \( r = 0.28011, p = 2.2204 \times 10^{-15} \)

No.3 canonical correlation coefficient: -0.1656
\[ u = -0.36991x_1 - 0.55525x_2 - 0.31421x_3 - 0.4035x_4 - 0.5416x_5 \]
\[ v = 0.17287y_1 + 0.15472y_2 - 0.10007y_3 + 0.95268y_4 - 0.07593y_5 - 0.13697y_6 - 0.06357y_7 \]
Linear regression between \( u \) and \( v \)
\[ u = 1.3933 \times 10^{-17} - 0.42811 \cdot v \]
Linear regression between \( v \) and \( u \)
\[ v = 3.3344 \times 10^{-17} - 0.064044 \cdot u \]
Pearson \( r = -0.16558, p = 3.7421 \times 10^{-6} \)

All canonical correlations and linear eigenmodels are statistically significant with \( p \leq 3.7421 \times 10^{-6} \).

### 3.3 Taste attribute class vs. chemical composition class

The calculated canonical correlations and linear eigenmodels for taste attribute class (Table 1) vs. chemical composition class (Table 3) are as follows (only the first three groups of results are given)

No.1 canonical correlation coefficient: -0.4127
\[ u = -0.4127 \cdot v \]

All canonical correlations and linear eigenmodels are statistically significant with \( p \leq 3.7421 \times 10^{-6} \).
\[
u_1 = -0.04686x_1 - 0.20019x_2 + 0.31934x_3 + 0.0961x_4 + 0.89134x_5 + 0.20428x_6 + 0.10154x_7
\]

\[
v_1 = 0.0215y_1 - 0.20678y_2 + 0.14947y_3 + 0.08131y_4 + 0.18111y_5 + 0.14266y_6 + 0.05774y_7 + 0.01542y_8 + 0.16934y_9 - 0.37881y_{10} + 0.0051y_{11} - 0.17602y_{12} + 0.10964y_{13} + 0.04404y_{14} + 0.28013y_{15} + 0.68449y_{16} + 0.04637y_{17}
\]

Linear regression between \(u_1\) and \(v_1\)

\[
u_1 = -6.2815 \times 10^{-18} - 0.35265v_1
\]

Linear regression between \(v_1\) and \(u_1\)

\[
v_1 = 9.4197 \times 10^{-18} - 0.48309u_1
\]

\[
p = \rho = 0.41275, p < 0.0001
\]

No.2 canonical correlation coefficient: 0.3378

\[
u_2 = 0.71443x_1 - 0.48748x_2 - 0.10732x_3 + 0.44241x_4 - 0.17876x_5 - 0.09808x_6 - 0.05594x_7
\]

\[
v_2 = 0.2426y_1 - 0.05867y_2 + 0.20696y_3 - 0.0215y_4 - 0.15208y_5 - 0.28724y_6 + 0.13741y_7 - 0.07541y_8 - 0.34951y_9 + 0.18074y_{10} + 0.00479y_{11} - 0.15872y_{12} - 0.15044y_{13} + 0.17092y_{14} - 0.18414y_{15} - 0.15651y_{16} - 0.30307y_{17} + 0.31954y_{18} + 0.04404y_{19} - 0.28013y_{20} + 0.68449y_{21} + 0.04637y_{22}
\]

Linear regression between \(u_2\) and \(v_2\)

\[
u_2 = -2.9681 \times 10^{-18} - 0.2791v_2
\]

Linear regression between \(v_2\) and \(u_2\)

\[
v_2 = -1.0703 \times 10^{-18} + 0.40893u_2
\]

Pearson \(\rho_1 = 0.33783, p = 0.00001\)

No.3 canonical correlation coefficient: -0.2617

\[
u_3 = 0.30342x_1 + 0.29661x_2 - 0.08946x_3 - 0.37942x_4 + 0.81177x_5 - 0.05866x_6 - 0.07469x_7
\]

\[
v_3 = -0.0796y_1 - 0.22354y_2 - 0.08376y_3 - 0.14587y_4 + 0.2728y_5 + 0.13554y_6 + 0.31414y_7 - 0.03226y_8 + 0.07441y_9 + 0.25989y_{10} - 0.25808y_{11} + 0.20452y_{12} - 0.02162y_{13} + 0.38821y_{14} - 0.19486y_{15} + 0.23667y_{16} + 0.16262y_{17} - 0.00395y_{18} + 0.43165y_{19} - 0.25989y_{20} - 0.25808y_{21} + 0.20452y_{22}
\]

Linear regression between \(u_3\) and \(v_3\)

\[
u_3 = -2.8082 \times 10^{-18} - 0.24437v_3
\]

Linear regression between \(v_3\) and \(u_3\)

\[
v_3 = 3.0321 \times 10^{-17} - 0.28017u_3
\]

Pearson \(\rho_3 = 0.33783, p = 1.4966 \times 10^{-13}\)

All canonical correlations and linear eigenmodels are statistically significant with \(p \leq 1.4966 \times 10^{-13}\).

### 3.4 Medicinal property class vs. chemical composition class

The calculated canonical correlations and linear eigenmodels for medicinal property class (Table 2) vs. chemical composition class (Table 3) are as follows (only the first three groups of results are given)

No.1 canonical correlation coefficient: -0.3723

\[
u_1 = -0.08775x_1 - 0.04354x_2 + 0.05099x_3 + 0.71101x_4 - 0.69452x_7
\]

\[
v_1 = -0.0989y_1 + 0.0559y_2 + 0.0374y_3 + 0.5289y_4 - 0.47068y_5 + 0.03507y_6 + 0.16535y_7 + 0.03438y_8 - 0.08874y_9 + 0.17874y_{10} - 0.00847y_{11} + 0.08475y_{12} + 0.05523y_{13} - 0.07855y_{14} - 0.24531y_{15} - 0.38085y_{16} + 0.04074y_{17} + 0.10758y_{18} + 0.2378y_{19} - 0.04105y_{20}
\]

Linear regression between \(u_1\) and \(v_1\)

\[
u_1 = -3.156 \times 10^{-17} - 0.40118v_1
\]

Linear regression between \(v_1\) and \(u_1\)

\[
v_1 = 1.6295 \times 10^{-17} - 0.34547u_1
\]

Pearson \(\rho_1 = 0.37228, p = 0.00001\)
No.2 canonical correlation coefficient: -0.2817
\[ u_2=0.40483x_1+0.36083x_2+0.30733x_3+0.35588x_4+0.69628x_5 \]
\[ v_2=-0.12512y_1+0.01661y_2-0.32913y_3+0.35257y_4+0.10112y_5+0.07457y_6+0.00013y_7-0.00566y_8+0.09068y_9+0.01763y_{10}+0.02882y_{11}-0.0068y_{12}+0.11194y_{13}+0.25656y_{14}-0.04017y_{15}+0.71883y_{16}+0.08554y_{17}-0.23867y_{18}-0.12695y_{19}+0.01813y_{20}-0.01006y_{21}+0.25932y_{22} \]
Linear regression between \( u_2 \) and \( v_2 \)
\[ u_2=1.0498*10^{-17}+0.99578v_2 \]
Linear regression between \( v_2 \) and \( u_2 \)
\[ v_2=3.8825*10^{-15}+0.079667u_2 \]
Pearson \( r=-0.28166, p=1.5543*10^{-15} \)

No.3 canonical correlation coefficient: 0.1889
\[ u_3=-0.2568x_1-0.34642x_2-0.50593x_3-0.30834x_4+0.68045x_5 \]
\[ v_3=0.1908y_1-0.01518y_2+0.07877y_3-0.49439y_4+0.10412y_5-0.06303y_6-0.0986y_7+0.05814y_8+0.22266y_9+0.11985y_{10}+0.07737y_{11}-0.13853y_{12}+0.04033y_{13}-0.38498y_{14}-0.05785y_{15}+0.48666y_{16}-0.0011y_{17}+0.19452y_{18}+0.01162y_{19}-0.18955y_{20}+0.35499y_{21}-0.07564y_{22} \]
Linear regression between \( u_3 \) and \( v_3 \)
\[ u_3=-1.3625*10^{-17}+0.37177v_3 \]
Linear regression between \( v_3 \) and \( u_3 \)
\[ v_3=2.4076*10^{-15}+0.095984u_3 \]
Pearson \( r=0.1889, p=1.2381*10^{-7} \)

All correlations and linear eignmodes are statistically significant with \( p\leq1.2381*10^{-7} \).

3.5 Meridians and colleterals class vs. medicinal function class
The calculated canonical correlations and linear eignmodes for meridians and colleterals class (Table 4) vs. medicinal function class (Table 5) are as follows (only the first three groups of results are given)

No.1 canonical correlation coefficient: 0.6934
\[ u_4=0.49467x_1-0.15557x_2+0.15316x_3+0.01776x_4-0.16345x_5-0.44071x_6-0.36465x_7+0.17153x_8+0.15786x_9+0.19911x_{10}+0.02969x_{11}+0.50847x_{12} \]
\[ v_4=0.13649y_1-0.90307y_2+0.13187y_3-0.29866y_4+0.00691y_5+0.03305y_6+0.04238y_7+0.00393y_8+0.05936y_9+0.0644y_{10}+0.03536y_{11}+0.07812y_{12}+0.0772y_{13}+0.01804y_{14}+0.03598y_{15}+0.25528y_{16}+0.12064y_{17}+0.24571y_{18}+0.07139y_{19}+0.04577y_{20}+0.13601y_{21}+0.15049y_{22}+0.07495y_{23}+0.18769y_{24}+0.00838y_{25}+0.04611y_{26}+0.11783y_{27}+0.01996y_{28}+0.13667y_{29}+0.05599y_{30}+0.03833y_{31}+0.01582y_{32}+0.04764y_{33}+0.09478y_{34}+0.07603y_{35}+0.00203y_{36}+0.09609y_{37}+0.0487y_{38}+0.0661y_{39}+0.000396y_{40}+0.18242y_{41}+0.00418y_{42}+0.07974y_{43}+0.11147y_{44} \]
Linear regression between \( u_4 \) and \( v_4 \)
\[ u_4=1.0577*10^{-17}+0.28372v_4 \]
Linear regression between \( v_4 \) and \( u_4 \)
\[ v_4=1.7112*10^{-17}+1.6948u_4 \]
Pearson \( r=0.69344, p=0 \)

No.2 canonical correlation coefficient: 0.642
\[ u_5=-0.02594x_1+0.34996x_2+0.13919x_3+0.7065x_4+0.47434x_5+0.25881x_6+0.0498x_7-0.11584x_8+0.15169x_9-0.02658x_{10}+0.14876x_{11}+0.06305x_{12} \]
IAEES www.iaees.org
Linear regression between $u_2$ and $v_2$

$$u_2 = 4.2091 \times 10^{-15} + 0.27965 \times v_2$$

Linear regression between $v_2$ and $u_2$

$$v_2 = -2.3996 \times 10^{-15} + 1.474 \times u_2$$

Pearson $r = 0.64203$, $p = 0$

No.3 canonical correlation coefficient: -0.5807

$$u_3 = 0.15424 x_1 + 0.2169 x_2 - 0.24008 x_3 + 0.45988 x_4 - 0.38404 x_5 + 0.1403 x_6 + 0.15062 x_7 + 0.22777 x_8 + 0.17801 x_9 + 0.44047 x_{10} - 0.41109 x_{11} + 0.16079 x_{12}$$

$$v_3 = -0.01495 y_1 + 0.09547 y_2 - 0.11743 y_3 - 0.12479 y_4 - 0.26804 y_5 + 0.18104 y_6 + 0.14225 y_7 - 0.02223 y_8 + 0.03199 y_9 + 0.1492 y_{10} + 0.14704 y_{11} + 0.12018 y_{12} + 0.14377 y_{13} + 0.06538 y_{14} - 0.1551 y_{15} - 0.1239 y_{16} + 0.00549 y_{17} - 0.07201 y_{18} - 0.12257 y_{19} - 0.06791 y_{20} - 0.07759 y_{21} - 0.10637 y_{22} - 0.20166 y_{23} - 0.09282 y_{24} + 0.04642 y_{25} - 0.08974 y_{26} - 0.29102 y_{27} + 0.20671 y_{28} - 0.11823 y_{29} + 0.10677 y_{30} - 0.08302 y_{31} + 0.04944 y_{32} - 0.08483 y_{33} + 0.00079 y_{34} - 0.05164 y_{35} - 0.02157 y_{36} + 0.01083 y_{37} - 0.0271 y_{38} + 0.04485 y_{39} + 0.01506 y_{40} + 0.08502 y_{41} + 0.3371 y_{42} + 0.4491 y_{43} + 0.07967 y_{44} + 0.03352 y_{45} + 0.05982 y_{46} + 0.03551 y_{47} + 0.03506 y_{48} + 0.00255 y_{49} - 0.4097 y_{50} + 0.00839 y_{51} - 0.10078 y_{52} - 0.03877 y_{53} + 0.01613 y_{54} + 0.03612 y_{55} + 0.11123 y_{56} + 0.02982 y_{57} + 0.08211 y_{58} + 0.01416 y_{59} + 0.09318 y_{60} + 0.09175 y_{61} + 0.05023 y_{62} + 0.0433 y_{63} + 0.03289 y_{64} + 0.1144 y_{65} + 0.2051 y_{66} + 0.1649 y_{67} + 0.0474 y_{68} + 0.0474 y_{69} - 0.01706 y_{70} - 0.00924 y_{71} - 0.00954 y_{72} - 0.0698 y_{73} - 0.00212 y_{74} + 0.00872 y_{75} + 0.05784 y_{76} + 0.31212 y_{77}$$

Linear regression between $u_1$ and $v_3$

$$u_3 = 5.5161 \times 10^{-16} + 0.24472 \times v_3$$

Linear regression between $v_3$ and $u_3$

$$v_3 = 3.3233 \times 10^{-11} + 1.3781 \times u_3$$

Pearson $r = -0.58073$, $p = 0$

All canonical correlations and linear eigmodels are statistically significant with $p \leq 10^{-15}$.

### 3.6 Taste attribute class vs. medicinal function class

The calculated canonical correlations and linear eigmodels for taste attribute class (Table 1) vs. medicinal function class (Table 5) are as follows (only the first three groups of results are given)

No.1 canonical correlation coefficient: -0.6338

$$u_1 = 0.28926 x_1 + 0.82418 x_2 + 0.31268 x_3 + 0.0237 x_4 + 0.0442 x_5 + 0.3203 x_6 - 0.18486 x_7$$

$$v_1 = -0.0332 y_1 - 0.26233 y_2 + 0.02572 y_3 - 0.08497 y_4 + 0.02954 y_5 - 0.04282 y_6 - 0.01083 y_7 + 0.0263 y_8 + 0.15673 y_9 - 0.07291 y_{10} + 0.08505 y_{11} + 0.04256 y_{12} - 0.09672 y_{13} + 0.31969 y_{14} + 0.01848 y_{15} - 0.05785 y_{16} - 0.05413 y_{17} + 0.0354 y_{18} - 0.12881 y_{19} - 0.02541 y_{20} + 0.0064 y_{21} + 0.25388 y_{22} + 0.07963 y_{23} - 0.00418 y_{24} + 0.01056 y_{25} + 0.23591 y_{26} - 0.09371 y_{27} + 0.08141 y_{28} + 0.0506 y_{29} + 0.05715 y_{30} + 0.03276 y_{31} - 0.15818 y_{32} + 0.0685 y_{33} + 0.0256 y_{34} - 0.01038 y_{35} - 0.15766 y_{36} + 0.02167 y_{37} + 0.0509 y_{38} + 0.07519 y_{39} + 0.11402 y_{40} + 0.15375 y_{41} + 0.20422 y_{42} + 0.03689 y_{43} - 0.05402 y_{44} - 0.00558 y_{45} + 0.01445 y_{46} + 0.00942 y_{47} + 0.15415 y_{48} + 0.1823 y_{49} - 0.03832 y_{50} + 0.12012 y_{51} + 0.04054 y_{52} + 0.2063 y_{53} - 0.08429 y_{54} + 0.04256 y_{55} - 0.00105 y_{56} - 0.04563 y_{57} - 0.05054 y_{58} + 0.05401 y_{59} + 0.18851 y_{60} + 0.03768 y_{61} + 0.06017 y_{62} + 0.07242 y_{63} + 0.01704 y_{64} + 0.16795 y_{65}$$

IAEES www.iaees.org
Pearson v3 Linear regression between Pearson v1 3257 v3 4 02 y77 y55 y1 3

All canonical correlations and linear eigimodels are statistically significant with \( p \leq 10^{-15} \).

### 3.7 Canonical correlation network of attribute classes

Similar to my previous research (Zhang, 2017b), I constructed the canonical correlation network of attribute

IAEES www.iaees.org
classes, as indicated in Fig. 1. In the network, only the greatest Pearson correlations are labeled, and the corresponding canonical correlation functions are listed in the sections above. All canonical correlations in the network are statistically significant ($p \leq 10^{-5}$). Different from Zhang (2017b), the canonical correlation network takes attribute classes as nodes and represents the correlations between attribute classes.

![Network Diagram](image.png)

**Fig. 1** Canonical correlation network of attribute classes.

**4 Discussion**
Similar to between-attribute correlation, the correlation between attribute classes may mostly tend to be the quasi-linear correlation, or nonlinear correlation (Zhang, 2012, 2015), as seen by the linear correlation coefficients above. On the other hand, taking into account extreme complexity of quasi-linear correlation or nonlinear correlation between attribute classes, linear correlation is a reasonable approximation to quasi-linear correlation and nonlinear correlation.

It should be noted that the canonical correlation function is an optimal linear combination of attributes (variables) to maximize the fitting goodness. Thus, they, as well as the derived linear eigenvectors from them, may be used as empirical models, but not as mechanism models. Its coefficient and sign are not suggested being used as importance or positive / negative interactions of attributes.

In a certain sense, the linear eigenvector serves more as a pattern of attributes of Chinese herbal medicines, rather than a predictive model. To predict the herbal attributes, other methods such as regression, discriminant analysis, etc (Qi, 2006; Liu et al., 2014) can be used.

**Acknowledgment**
We are thankful to the support of Discovery and Crucial Node Analysis of Important Biological and Social Networks (2015.6-2020.6), from Yangling Institute of Modern Agricultural Standardization, China.
References