راهنمای کار با نرم‌افزار

VERSION 2.1

۱- مقدمه

۱-۱- زمینه‌های کاربرد

$NTSYS_{PC}$ نرم‌افزاری است که برای یافتن و نشان دادن ساختار اطلاعات چندمتغیره استفاده می‌شود. این نرم‌افزار معمولاً در علوم بیولوژی مورد استفاده قرار می‌گیرد. اما این برنامه به‌طور گسترده در علوم مورفومتریک، اکولوژی، علوم طبیعی، مهندسی و انسان شناسی استفاده می‌گردد. اصطلاحات mathematical taxonomy و automatic classification برای تشریح این زمینه از کاربردها استفاده می‌شود.

در بخش بیولوژی سیستماتیک، یک فرد را می‌توان از دو روش کلاس‌بندی، متمایز کرد. **فنتیک‌های** یک فرد که به کشف و تشریح الگوهای تنوع بیولوژی و شکل‌گیری کلاس‌بندی برپایه همه شباهت‌های محاسبه شده از اطلاعات چندمتغیره بستگی دارد. این اطلاعات معمولاً در مطالعات مورفومتریک استفاده می‌شود. Cladisticهای یک فرد که در تاریخچه تکاملی ارگانیسم تحت مطالعه استفاده می‌شود و به عنوان یک مبنا برای کلاس‌بندی استفاده می‌شود.

روش‌های موجود در NTSYSpc به طور زیادی با زمینه فنتیکی ارتباط دارند. اگرچه این‌ها تفسیر بهتری برای اطلاعات چندمتغیره می‌کنند. برنامه‌های دیگری جهت روش‌های فیلوژنی وجود دارد.

۱-۲- واحدهای برنامه در NTSYSpc:

در زیر تعریف‌های مختصری از واحدهای محاسبه‌ای در NTSYSpc آورده شده است. NTSYSpc فقط محدود به آنالیزهایی که در زیر ذکر شده نمی‌شود. این واحدها را می‌توان برای ساختن انواع آنالیزهای مختلف دیگر استفاده کرد (مثلاً آنالیز مختصات اصلی Gower را می‌توان بوسیله واحدهای SIMINT، DCENTER و EIGEN انجام داد).

CANPLS: انجام دادن همبستگی استاندارد و آنالیزهای two-block partial least-squares.

CONSENSUS: محاسبه یک درخت عمومی برای دوتا از دو یا بیشتر درخت (از قبیل ارتباط چندتایی درخت‌های SAHN یا بین دو روش مختلف). چندین شاخص عمومی همچنین برای اندازه‌گیری درجه مطابقت بین درخت‌ها محاسبه می‌کند.

COPH: یک ماتریس مقدار کوفنتیک (ماتریس مقادیر التراماتریک) از یک ماتریس درختی تولید می‌کند. همچنین می‌تواند یک ماتریس از فواصل مسیر طولی از نتایج برنامه NJOIN محاسبه کند.

CORRESP: آنالیز ارتباط.

CPCA : آنالیز مولفه‌های اصلی. تلاش می‌کند جهت برازش یک تک مجموعه از وکتورهای ویژه برای یک سری از ماتریس‌های واریانس-کوواریانس.

CVA: آنالیز بردارهای متداول را انجام می‌دهد.

DCENTER: یک double-centering از ماتریس‌های تشابه و عدم تشابه را انجام می‌دهد. ماتریس به دست آمده می‌تواند عاملی شود جهت انجام آنالیز مختصات اصلی.

EIGEN: محاسبه ماتریس‌های بردارهای ویژه . مقادیر ویژه از یک ماتریس تشابه سیمتریک موجود.

FOURIER: محاسبه تعییرات چهارتایی و چهارتایی بیضوی.

FOURPLOT: ترسیم نمودار و برآورد کردن نمودارهای ترسیم شده توسط مدل FOURIER.

MDSCALE: آنالیز مقیاس چندبعدی خطی و غیر متریک. این می‌تواند به عنوان یک جانشین برای آنالیز مختصات اصلی استفاده شود.

MOD3D: ترسیم یک دیاگرام پراکندگی سه طرفه به عنوان یک شکل سه بعدی.

MST: محاسبه حداقل طول درخت کشیده شده از ماتریس تشابه و عدم تشابه.

MULREG: انجام انواع مختلفی از آنالیزهای رگرسیون.

MXCOMP: محاسبه دو ماتریس سیمتریک بوسیله محاسبه همبستگی این دو ماتریس و سپس ترسیم یک دیاگرام پراکندگی.

MXPLOT: ترسیم یک دیاگرام پراکندگی دوبعدی از ردیف‌ها یا ستون‌های یک ماتریس.

OUTPUT: ماتریس‌ها را برای پرینت درون صفحات قرار می‌دهد.

PLOT: نمودار یک یا چند ستون یک ماتریس برضد یک ستون انتخاب شده.

POOLVCV: محاسبه یک ماتریس واریانس-کوواریانس درون گروهی مشترک از اطلاعات دو یا چند ماتریس.

SAHN: انجام دادن روش‌های متوالی، پیوندی، طبقاتی و خوشه‌بندی آشیانه‌ای.

SIMGEND: محاسبه ماتریس‌های ضرایب فواصل ژنتیکی از فراوانی ژنی و اطلاعات توالی DNA.

SIMINT: محاسبه شاخص‌های مختلف تشابه و عدم تشابه.

SIMQUAL: محاسبه ضرایب مختلف جمعیت برای اطلاعات کیفی.

STAND: انجام تغییرات خطی از اطلاعات ماتریس.

TRANSF: محاسبه تغییرات خطی و غیرخطی از ردیف‌ها یا ستون‌های یک ماتریس.

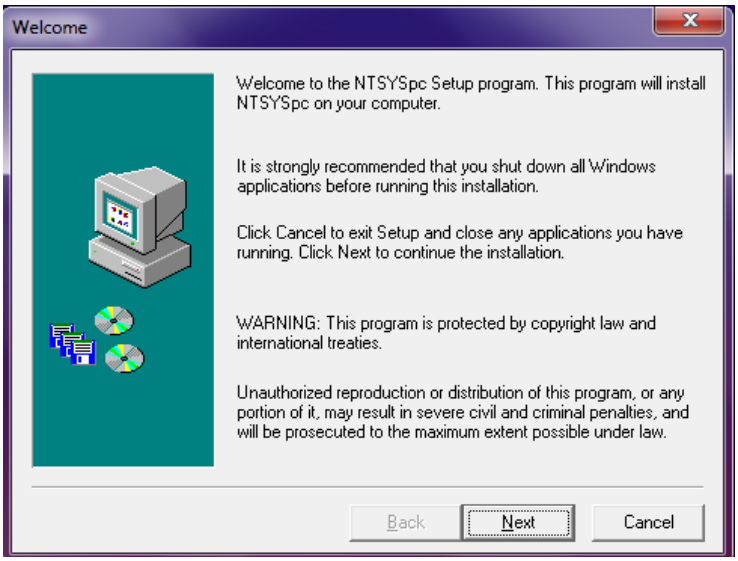
TREE: نمایش درخت فنتیکی و فیلوژنتیکی.

۱-۳- چگونگی شروع کار با نرم افزار

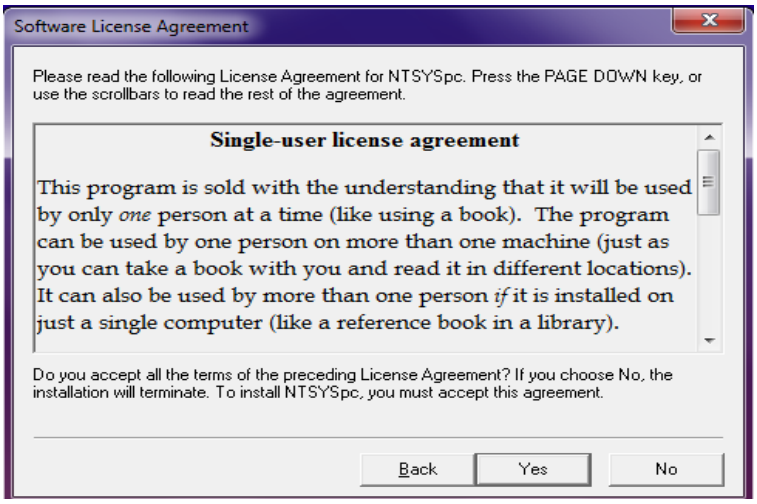
Installation : ابتدا باید نرم افزار را نصب کنیم. این عمل به راحتی و خیلی سریع انجام پذیر است. اول فایل مربوط به Setup را انتخاب می‌کنیم. سپس همانطور که در شکل۱ دیده می‌شود از کاربر پسورد خواسته می‌شود که این پسورد در فایل مربوط به پسورد موجود است.

شکل۱

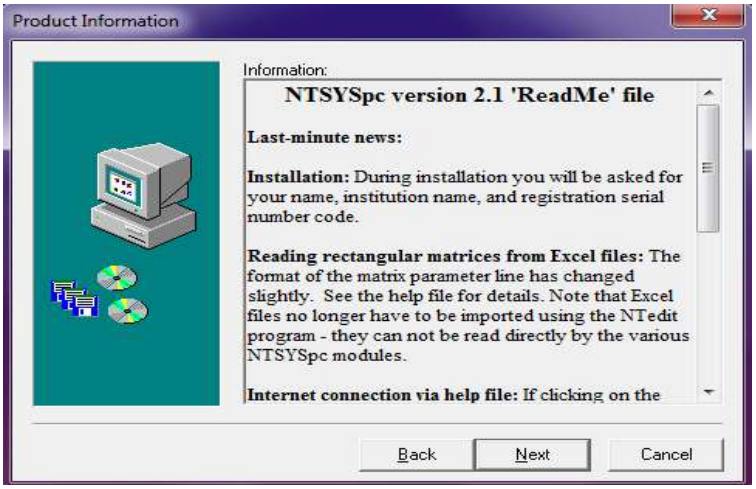
سپس کلیه مراحل نصب را همانند شکل‌های زیر پیش می‌بریم.



شکل۲

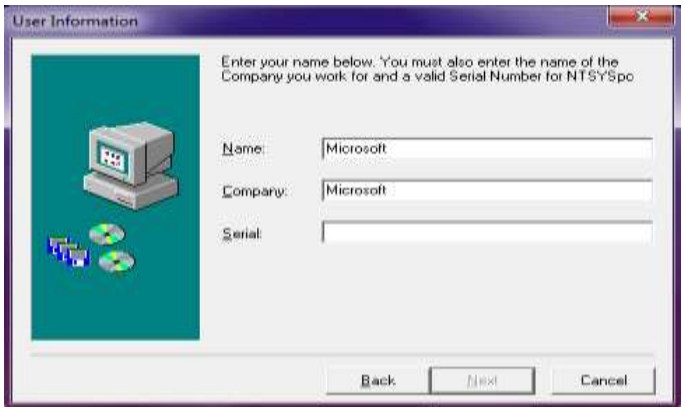


شکل ۳

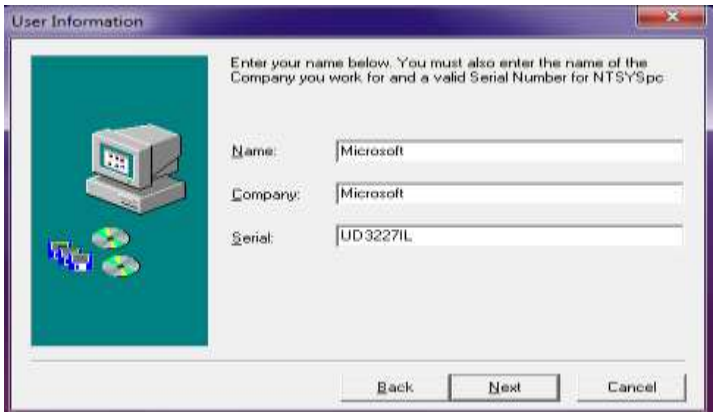


شکل ٤

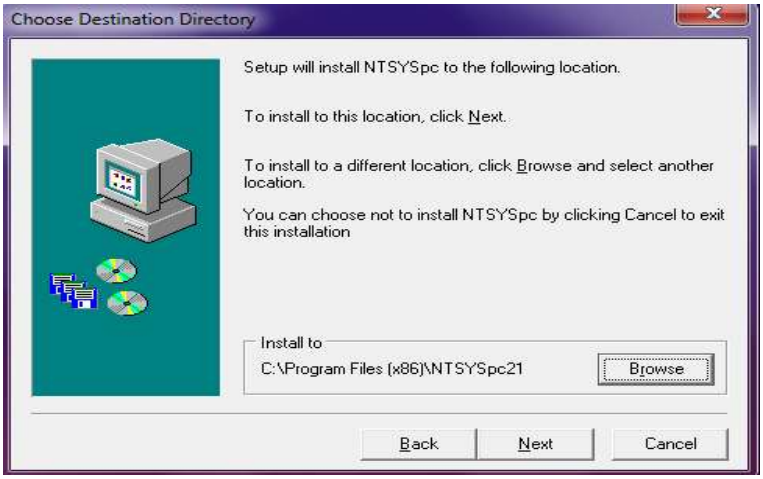
همانطور که در شکل ۵ دیده می‌شود از کاربر سریال خواسته شده است که این سریال در فایلی که پسورد قرار داشت موجود است.



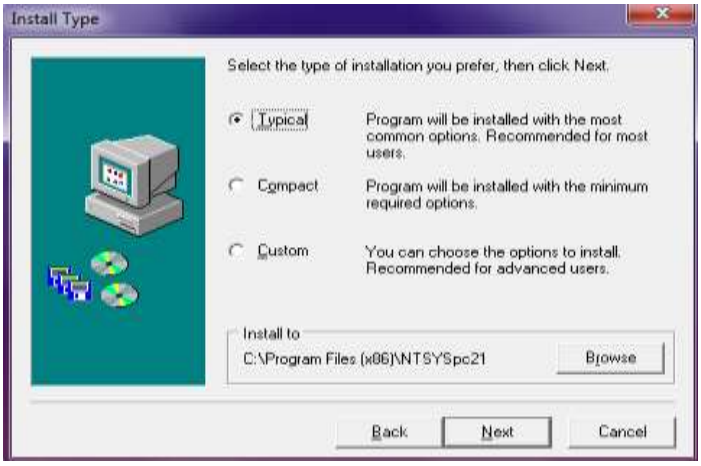
شکل ۵



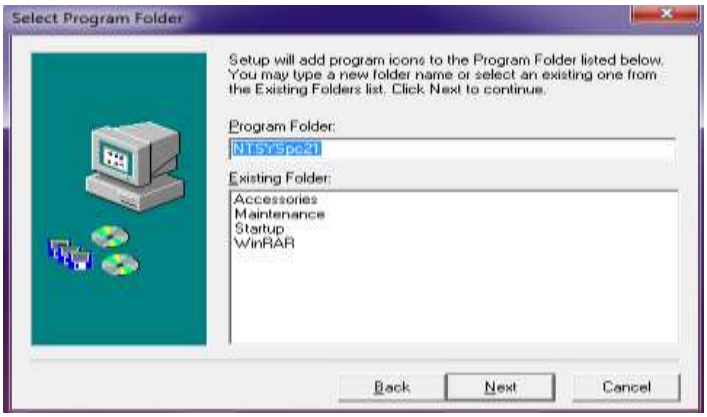
شکل٦



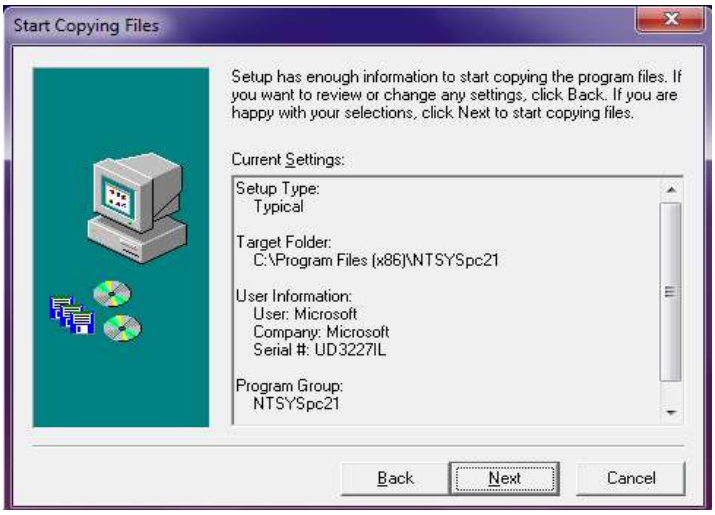
شکل٧



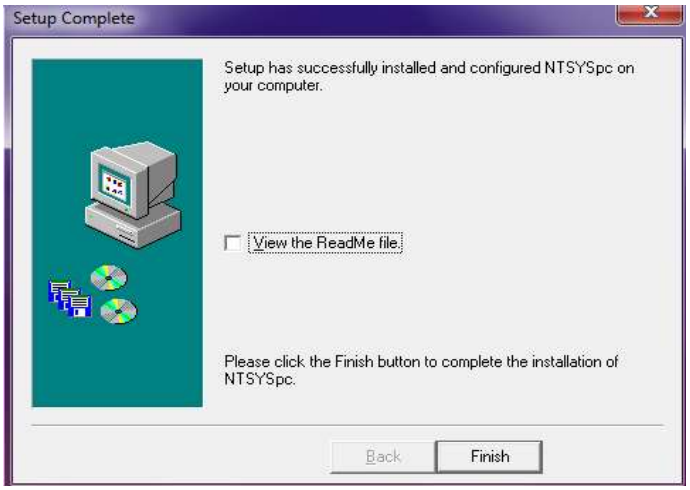
شکل٨



شکل ۹



شکل ۱۰

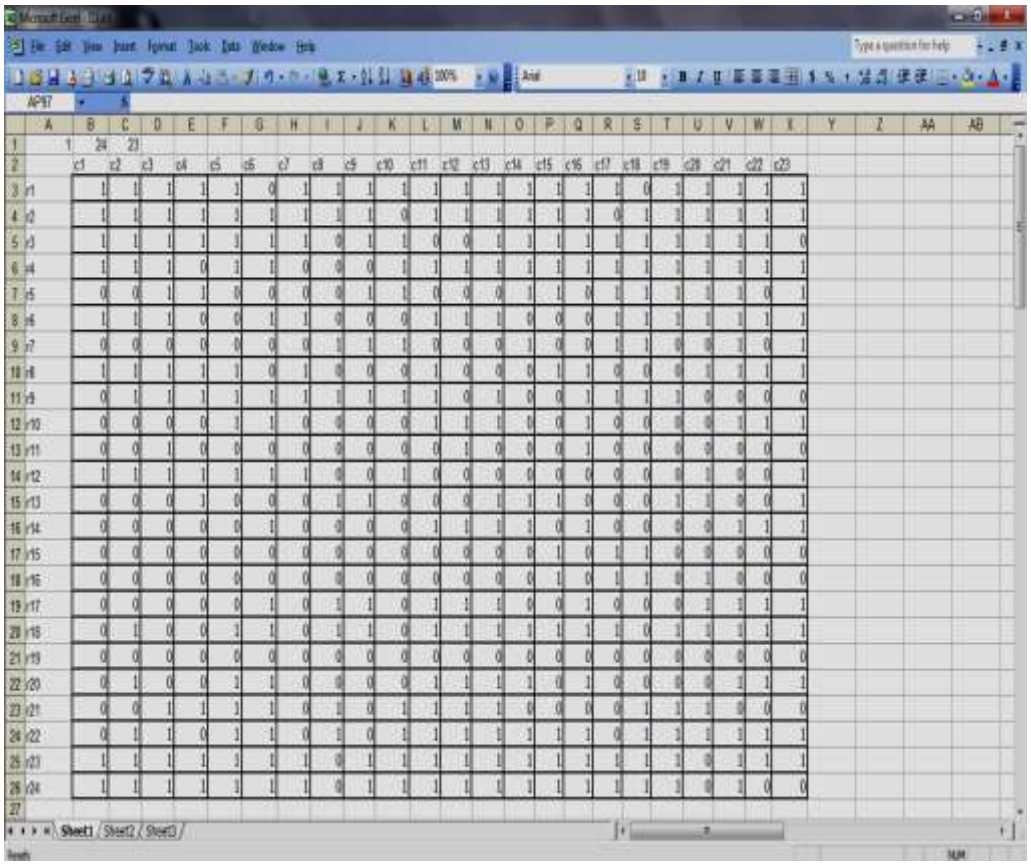


شکل ۱۱

پس از کامل شدن مراحل بالا، نرم‌افزار به همراه یک نرم‌افزار ویرایشی کمکی (NTedit) نصب می‌شود.

۲- استفاده از این نرم‌افزار در آنالیزهای مرتبط با مطالعات تنوعات ژنتیکی بوسیله نشانگرهای مولکولی

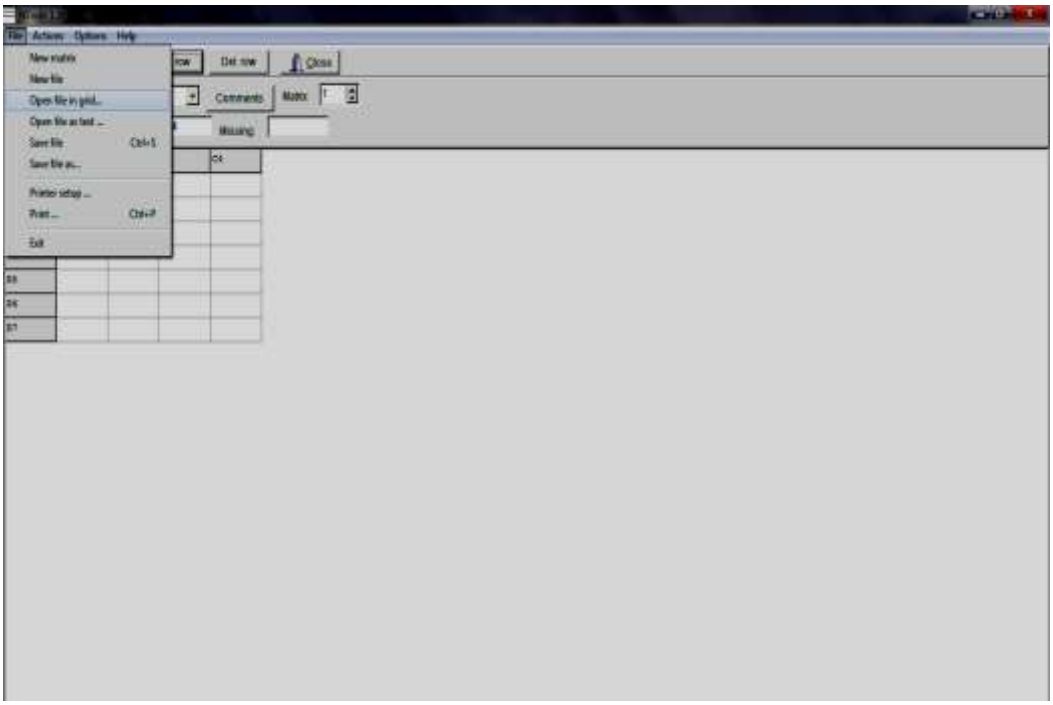
ابتدا باید در excel کدهای صفر و یک را وارد کنیم.



شکل۱۲

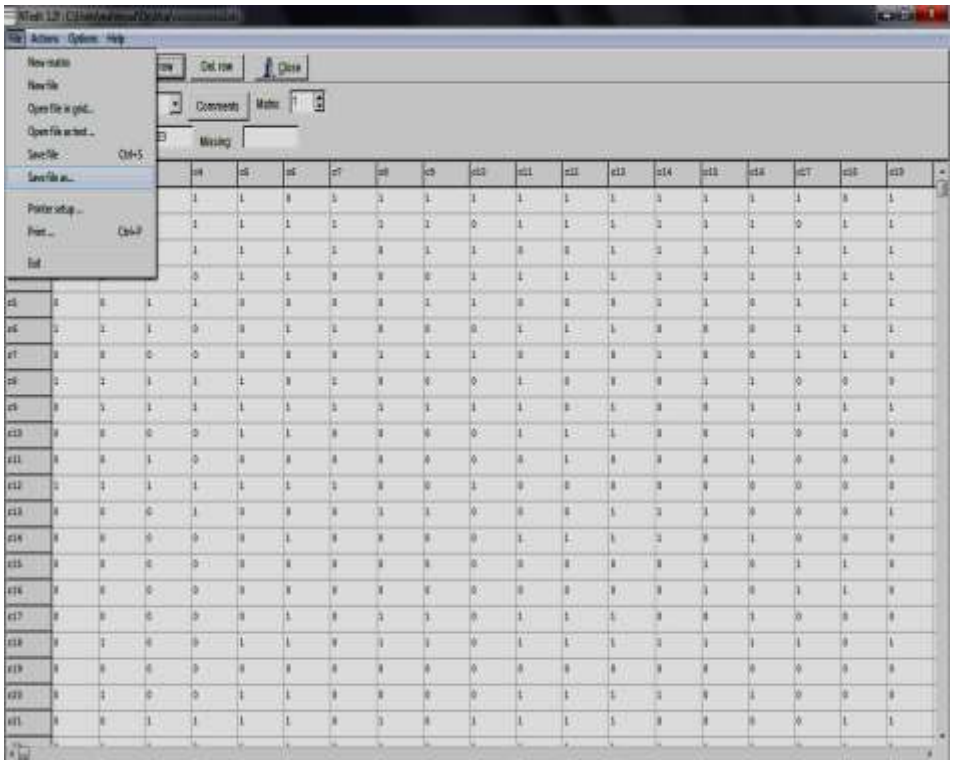
در شکل بالا A1 به طور معمول عدد ۱ را می‌گذاریم که بیانگر نام ماتریس است. B1 تعداد باندهای پلی‌مورف است. باید توجه کرد که باندهای مونومرف را حذف کنیم و اگر این کار را انجام ندادیم این باندها را باید در تعداد باندهای پلی‌مورف محاسبه کنیم.

C1 تعداد نمونه‌های مورد بررسی را نشان می‌دهد. ردیف ۲ شامل نام نمونه‌ها می‌باشد. ستون A شامل تعداد باندها می‌باشد. پس از کامل شدن فایل مورد نظر را save می‌کنیم. سپس وارد نرم افزار NTedit می‌شویم و فایل اکسل را وارد این نرم افزار می‌کنیم.



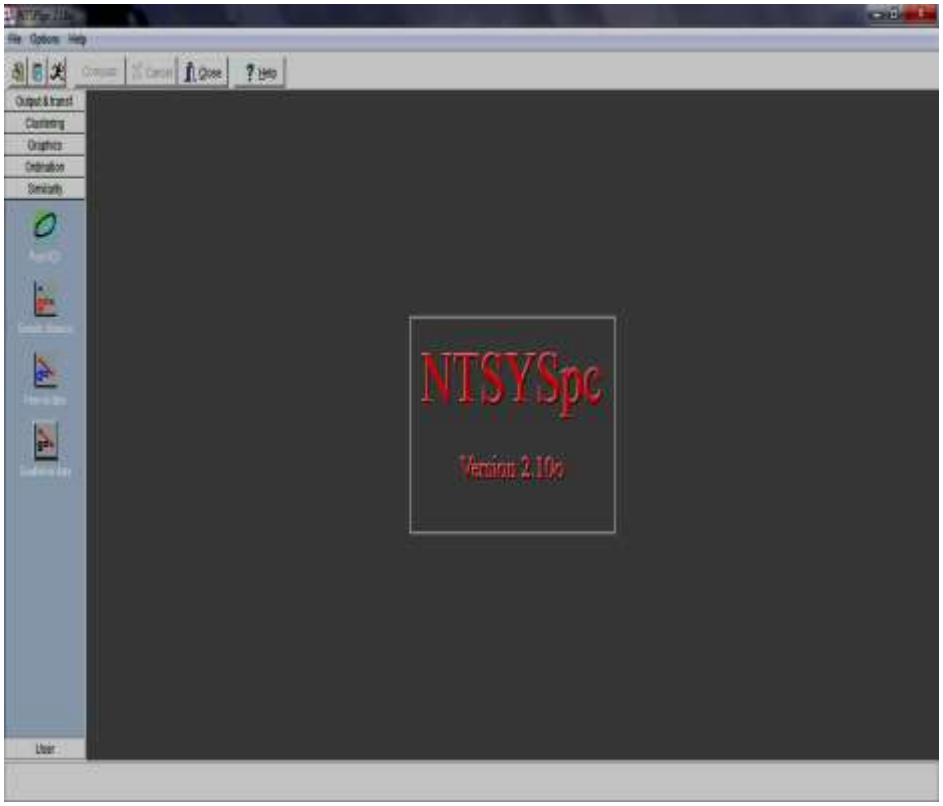
شکل۱۳

سپس باید این فایل را ذخیره کنیم.

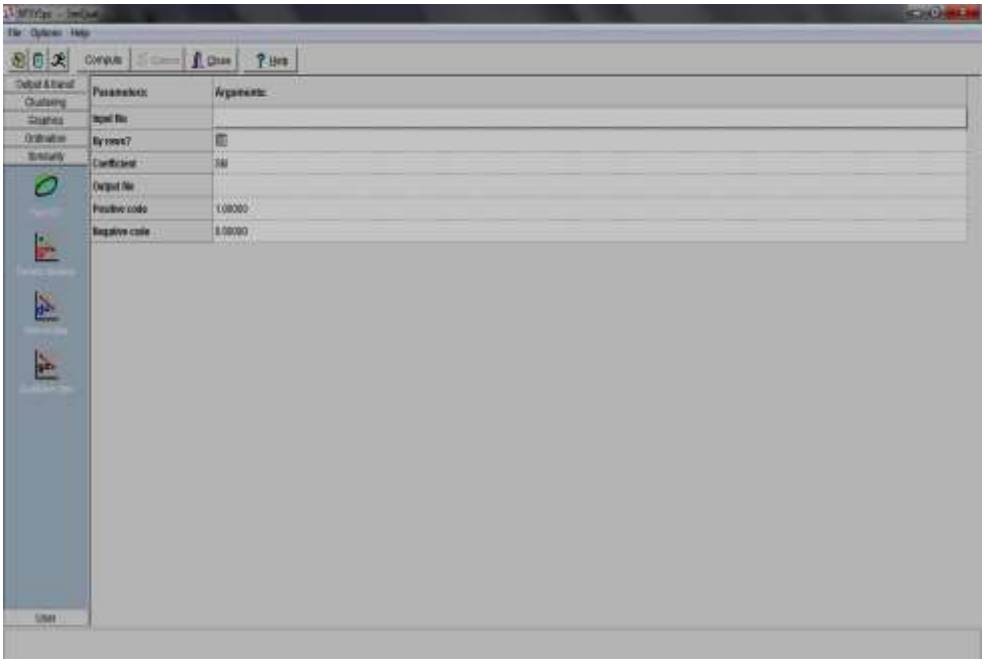


شکل۱٤

در ادامه باید وارد خود برنامه اصلی NTSYS شویم.



شکل۱۵

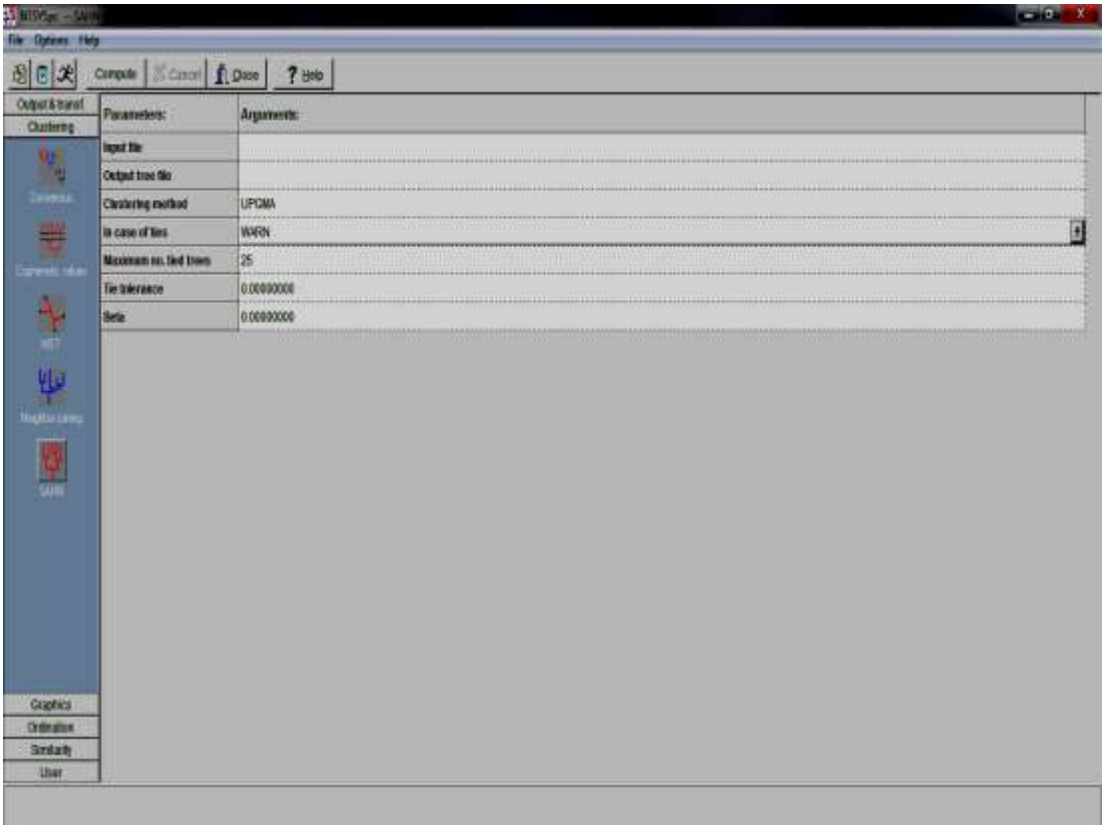


شکل۱۶

در قسمت Input file باید فایل NTedit ساخته شده را وارد کنیم. در قسمت Output file آدرس ماتریس مورد نظر را که می‌خواهیم بسازیم وارد می‌کنیم. این ماتریس همان **ماتریس تشابه** یا **فاصله**

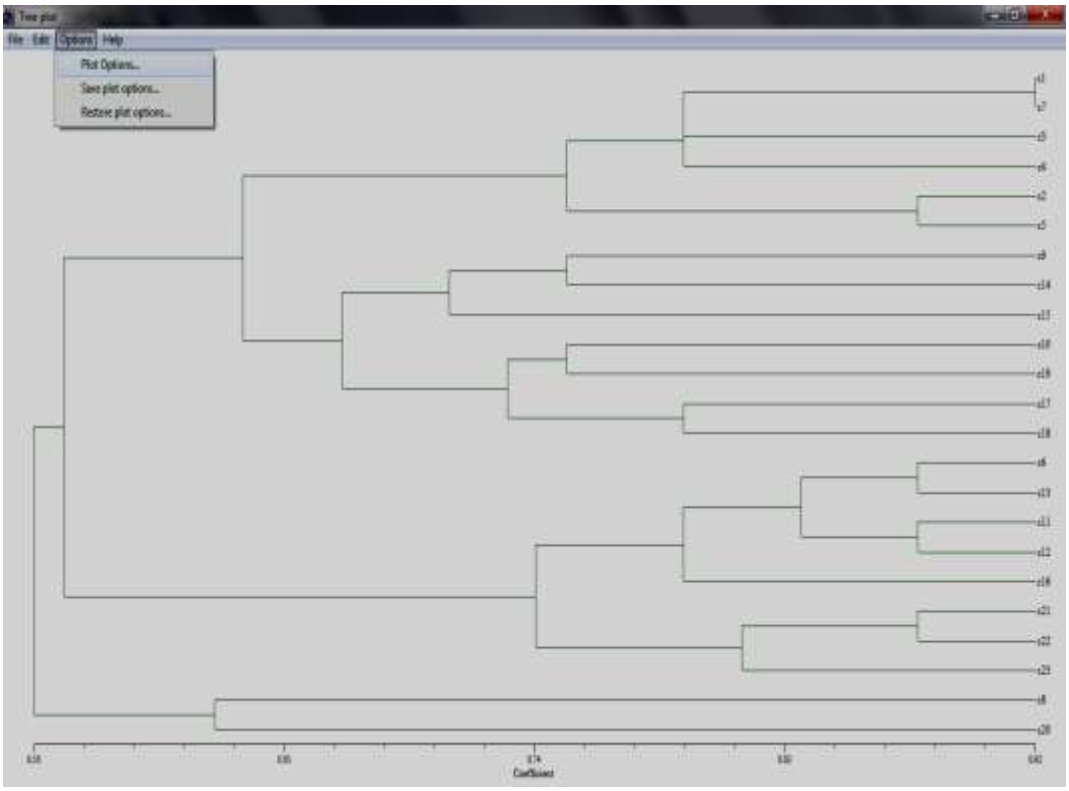
می‌باشد. در قسمت Coefficient ۱۶ روش مختلف وجود دارد که بنا به کارایی مورد نظر انتخاب می‌شود. البته معمولاً برای نشانگرهای مولکولی سه روش simple matching (SM) ، Jacard (J) و Dice انتخاب می‌شود. در آخر باید جهت محاسبه گزینه compute را انتخاب کنیم.

سپس باید اقدام به ترسیم دندروگرام کنیم. برای این منظور همانند شکل ۱۷ وارد بخش clustering می‌شویم. بخش SAHN را انتخاب می‌کنیم.

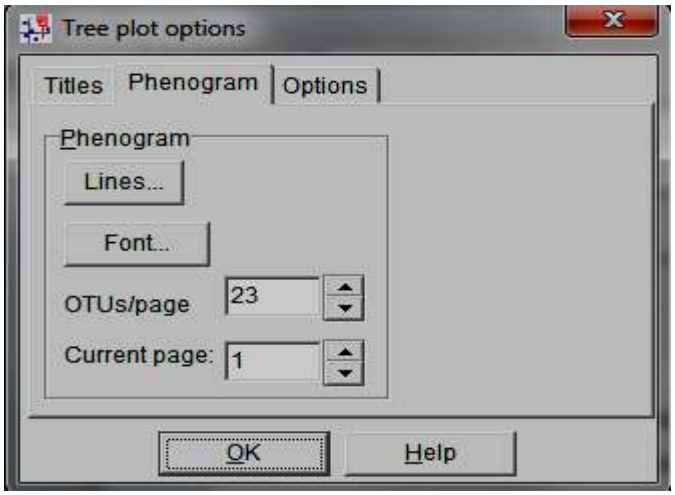


شکل۱۷

در قسمت Input file باید ماتریس تشابه ساخته شده را وارد کنیم و در قسمت Output tree file باید یک آدرس برای ماتریس درختی مشخص کنیم. در قسمت clustering method روش کلاستربندی را انتخاب می‌کنیم. ۸ روش وجود دارد که معمولاً سه روش complete، UPGMA و Single انتخاب می‌شود. جهت محاسبه گزینه compute را انتخاب کنیم. با ترکیب این روش‌ها ۹ دندروگرام به دست می‌آید که با نظر محقق بهترین آن‌ها باید انتخاب شود. در شکل ۱۸ و ۱۹ نشان داده شده که چگونه تمام نمونه‌ها در یک صفحه نشان داده شود (اگر تعداد نمونه‌ها بیش از ۳۰تا بود).



شکل۱۸

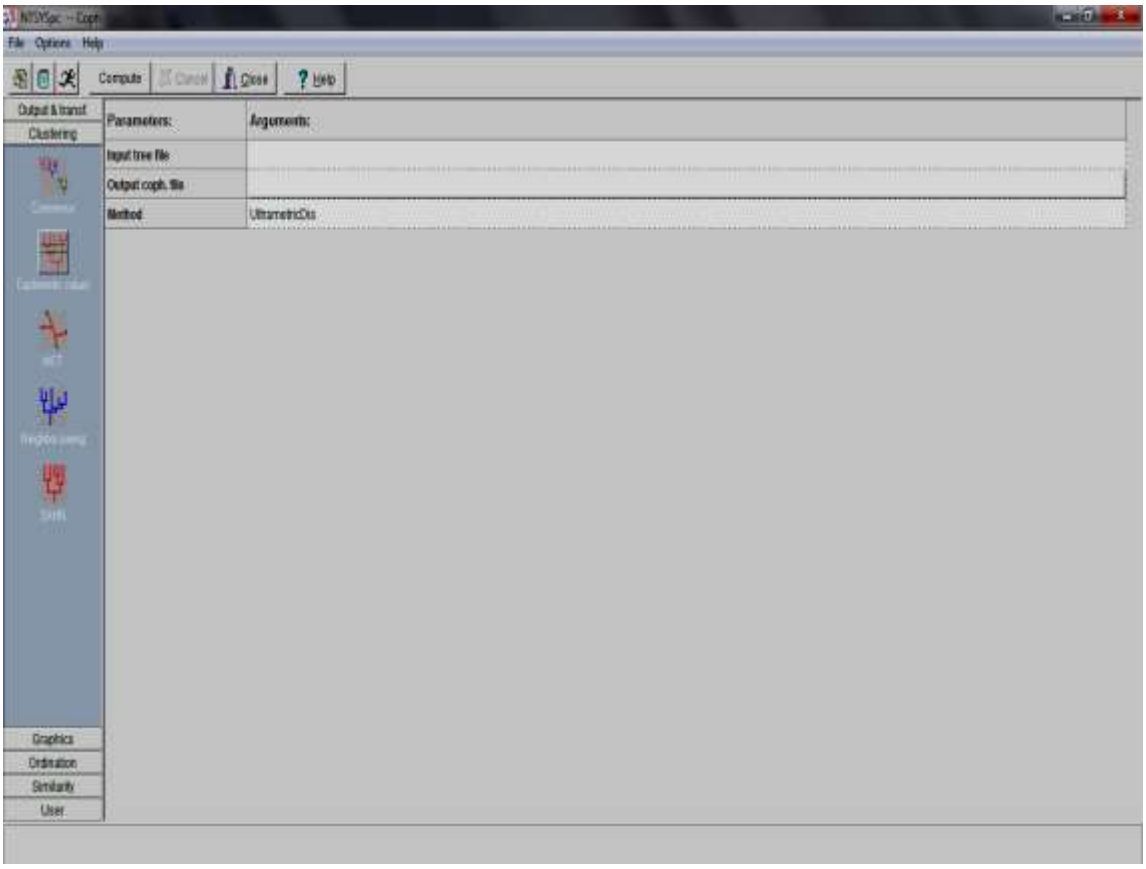


شکل۱۹

در قسمت OTUs/page تعداد کل نمونه‌ها را بوسیله فلش‌های بالا یا پایین تعیین می‌کنیم که در این مثال ۲۳ عدد می‌باشد.

یکی از راه‌های انتخاب بهترین دندروگرام به دست آوردن ضریب کوفنتیک است که اگر این عدد ۰/۹ یا بیشتر باشد نشاندهنده دندروگرام مطلوب است. البته این قائده همیشه صادق نیست. برای این کار در

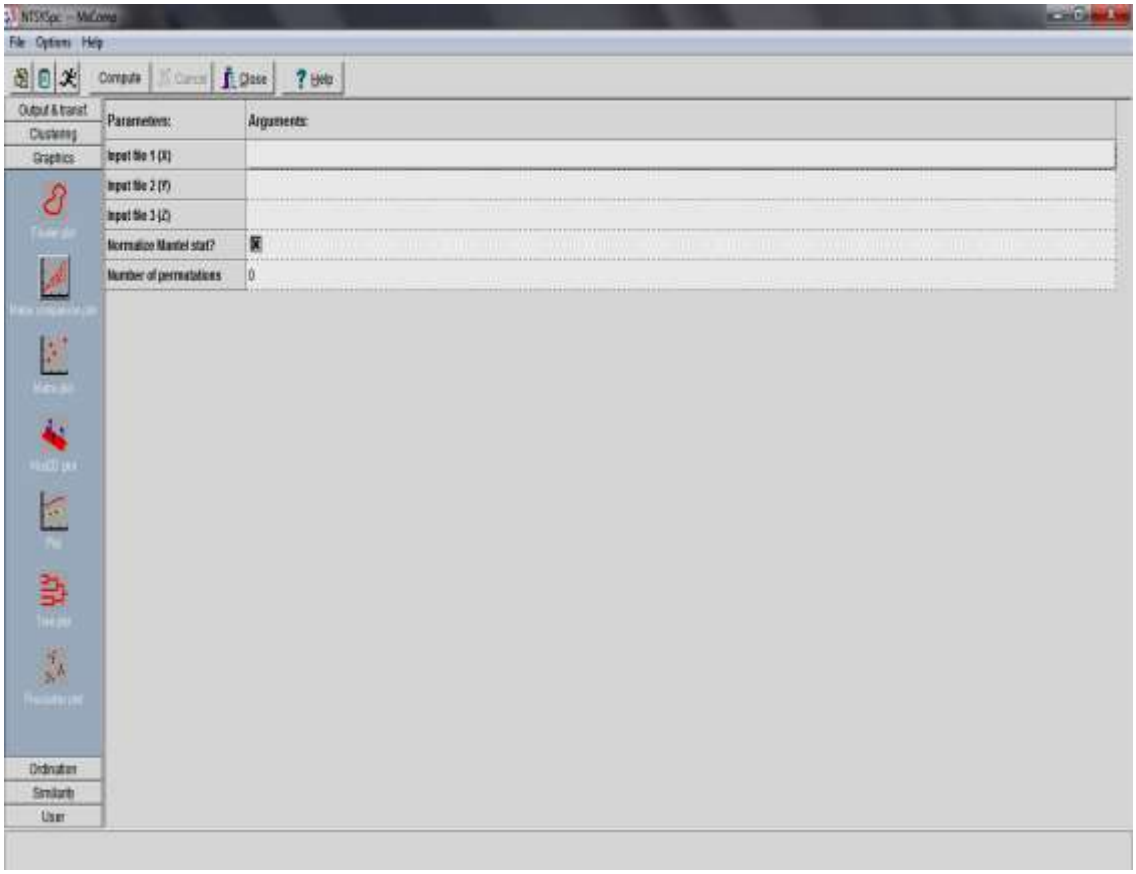
همان قسمت clustering بخش Cophenetic values را انتخاب می‌کنیم تا ماتریس کوفنتیک را بسازیم.



شکل ۲۰

در بخش Input tree file ماتریس درختی که در بخش قبل ساخته شده بود را وارد می‌کنیم. در قسمت Output coph. File آدرس ماتریس کوفنتیک مورد نظر را وارد می‌کنیم. جهت محاسبه گزینه compute را انتخاب کنیم.

سپس وارد بخش Graphics می‌شویم و گزینه Matrix comparison plot را انتخاب می‌کنیم.



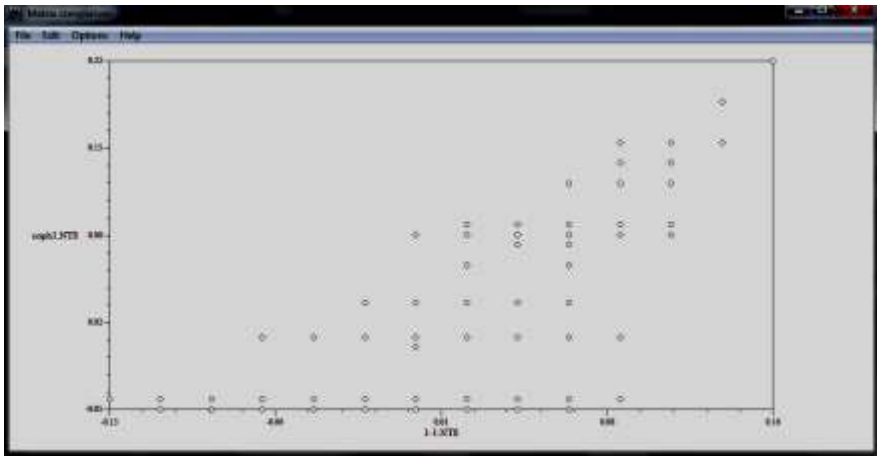
شکل ۲۱

در بخش Input file 1 (X) ماتریس تشابه‌ای که در قسمت اول ساختیم وارد می‌کنیم.

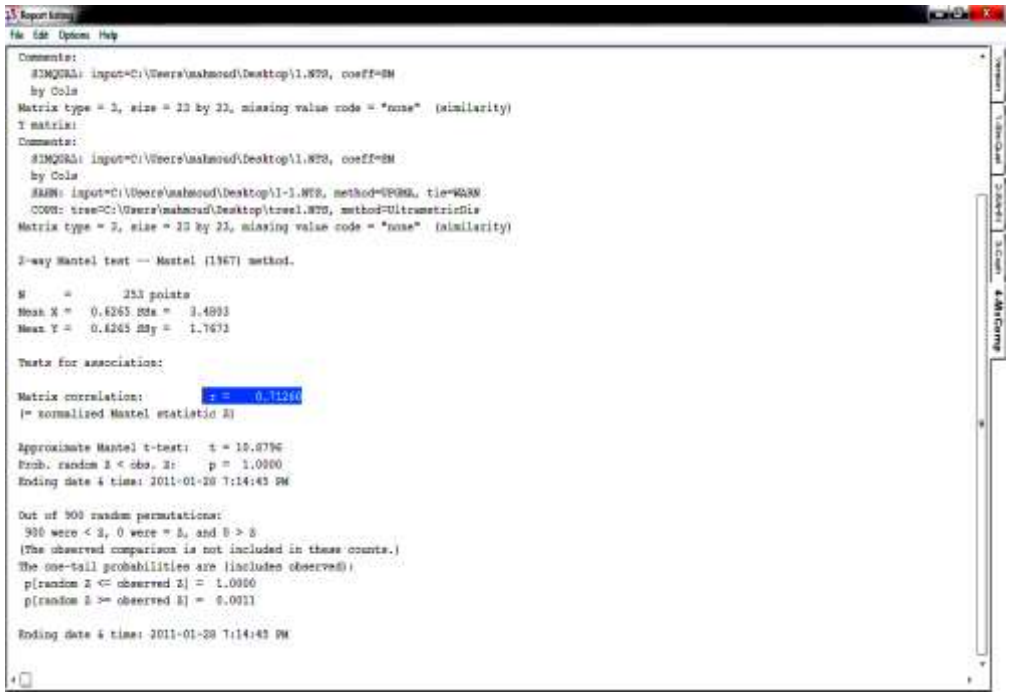
در بخش Input file 2 (Y) ماتریس کوفنتیک که در قسمت قبل ساختیم وارد می‌کنیم.

در بخش Number of permutations از عدد ۱ تا ۹۹۹ را می‌توانیم انتخاب کنیم. این عدد نشان دهنده تعداد دفعاتی است که این دو ماتریس وارد شده را با هم مقایسه می‌کند. معمولاً عدد ۹۰۰ انتخاب می‌شود. جهت محاسبه گزینه compute را انتخاب کنیم.

نتیجه یک نمودار پراکندگی و یک لیست گزارش می‌باشد که در مورد مثال آورده شده در شکل‌های زیر موجود است.



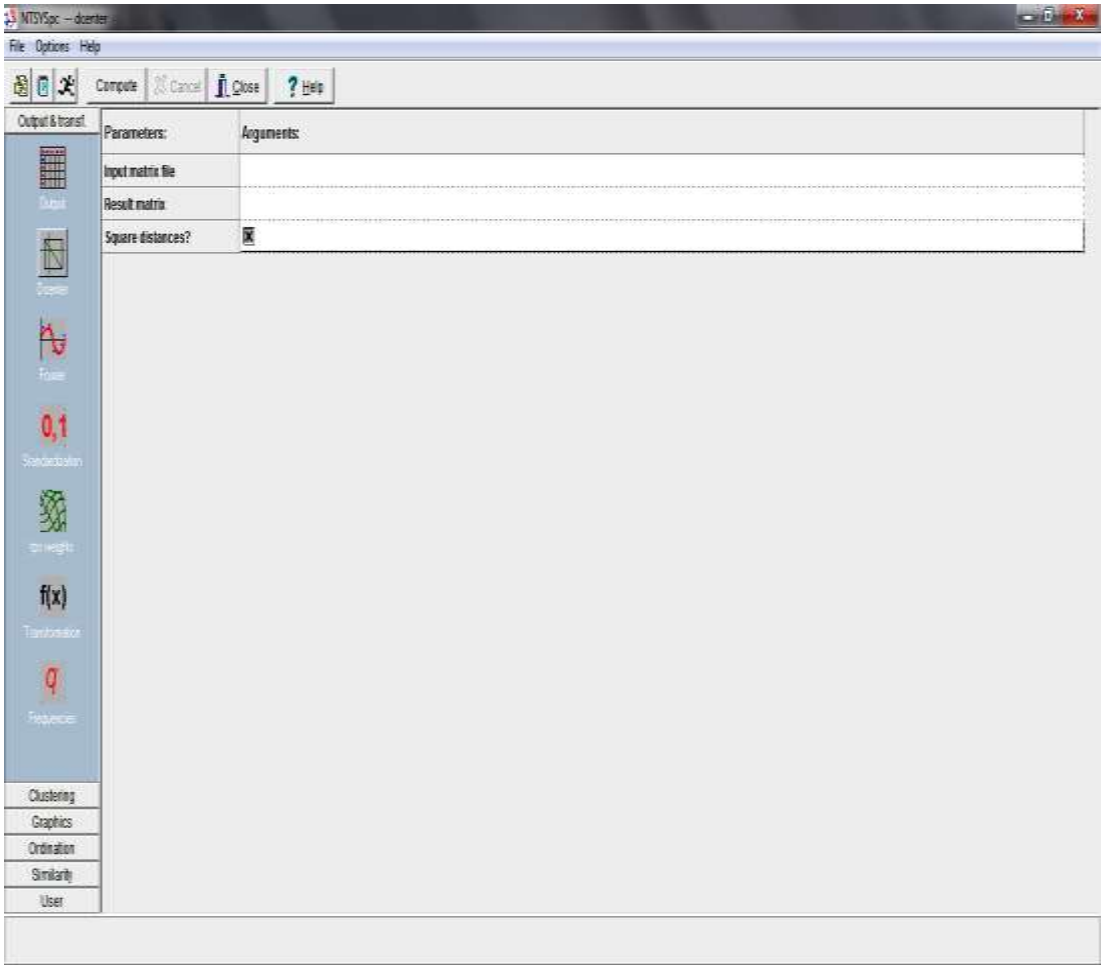
شکل۲۲



شکل۲۳

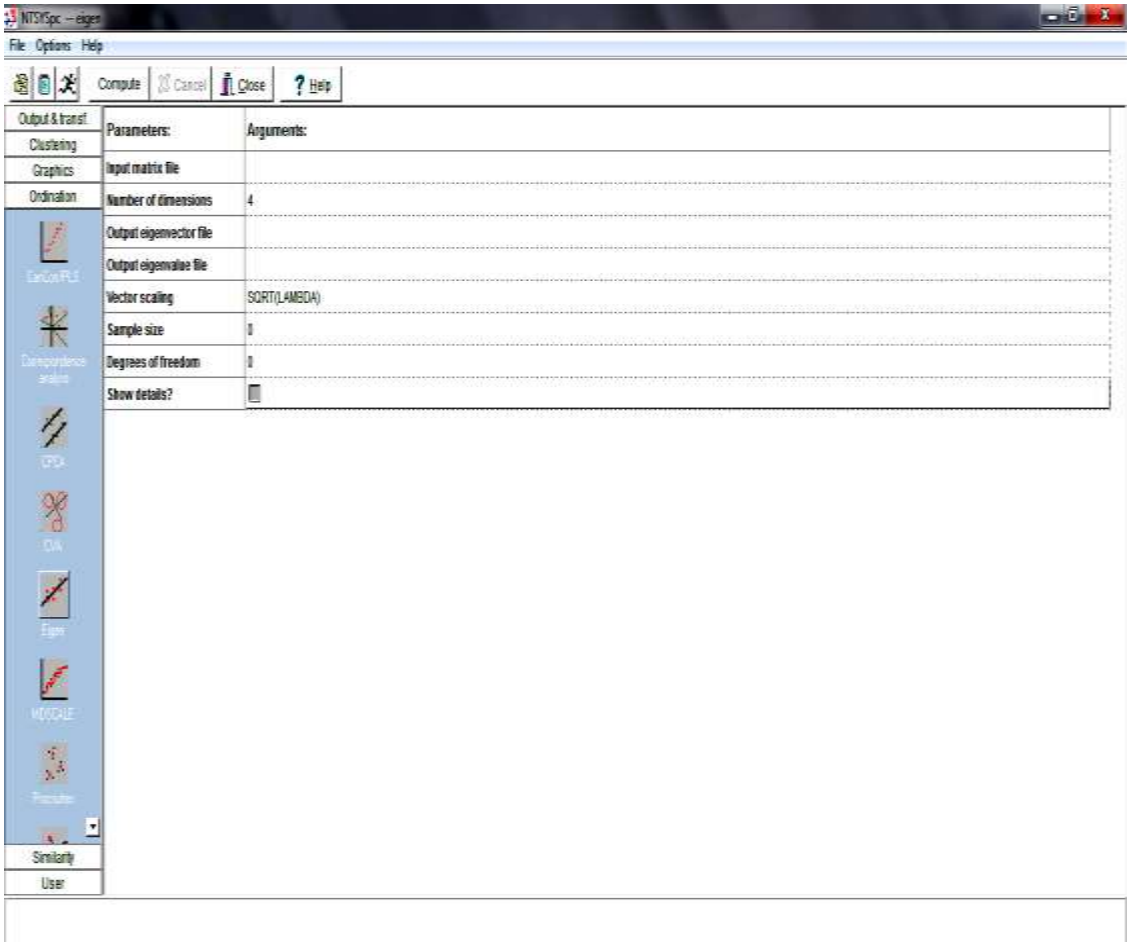
عدد r که در شکل نشان داده شده است همان ضریب کوفنتیک می‌باشد.

در مرحله بعدی بایستی تجزیه به مولفه‌های اصلی را انجام دهیم. ابتدا باید وارد بخش Output&transf. می‌شویم و گزینه Dcenter را انتخاب می‌کنیم.



شکل ۲٤

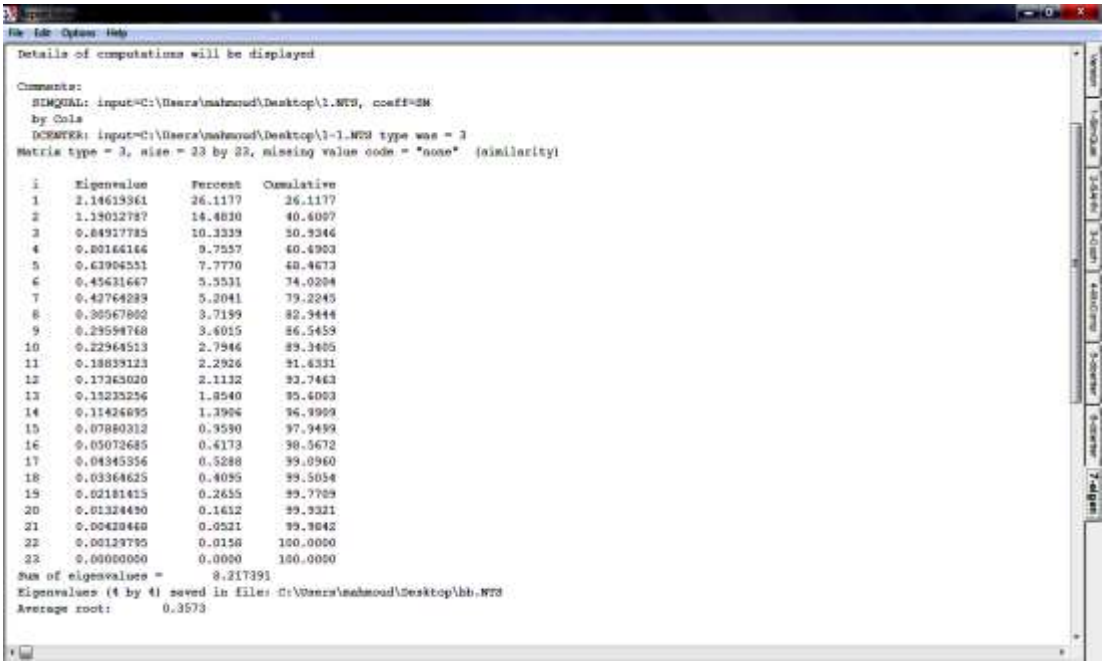
در بخش Input matrix file همان ماتریس تشابه اولیه را وارد می‌کنیم. در بخش Result matrix آدرس ماتریسی که می‌خواهیم بسازیم را وارد می‌کنیم. جهت محاسبه گزینه compute را انتخاب کنیم. در ادامه باید مقادیر ویژه را به دست آوریم. بنابراین وارد بخش ordination می‌شویم و گزینه Eigen را انتخاب می‌کنیم.



شکل ۲۵

در قسمت Input matrix file ماتریسی را که در بخش قبل ساختیم وارد می‌کنیم. در بخش Output eigenvalue file آدرس فایل مورد نظر را که می‌خواهیم بسازیم وارد می‌کنیم. جهت محاسبه گزینه compute را انتخاب کنیم.

یک لیست از گزارشات به دست می‌آید که شامل مقادیر ویژه (شکل۲۶) و واریانس مورد انتظار است (شکل۲۷).



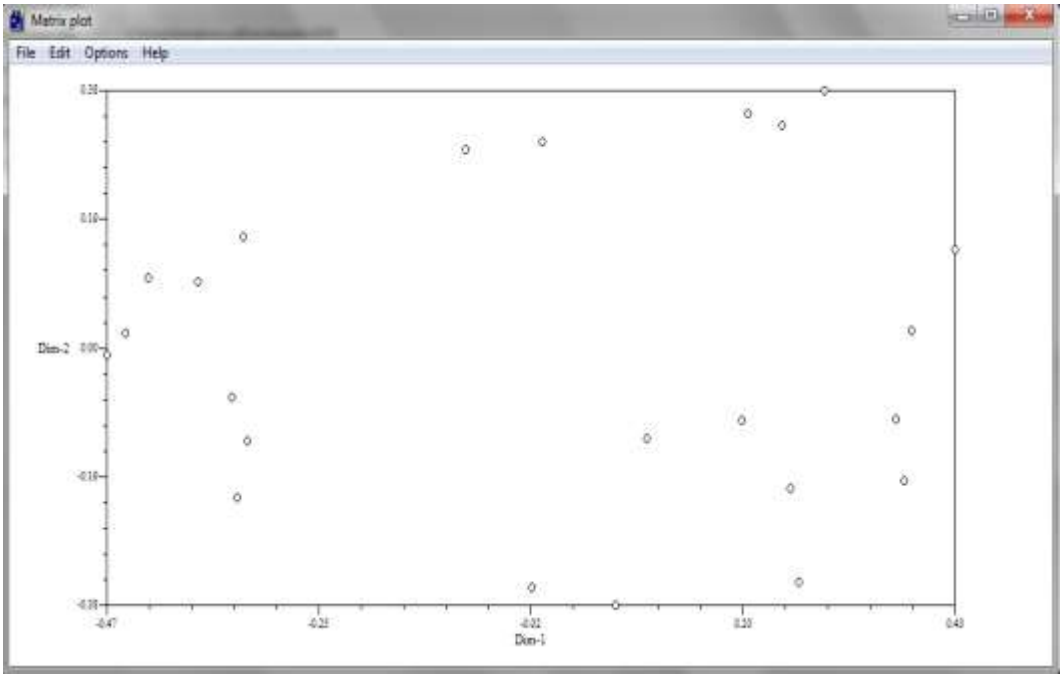
```
File  Edit  Options  Help
Details of computations will be displayed

Comments:
  SIMQUAL: input=C:\Users\mahmoud\Desktop\1.NTS, coeff=SM
  by Cola
  DCREATE8: input=C:\Users\mahmoud\Desktop\1-1.NTS type was = 3
Matrix type = 3, size = 23 by 23, missing value code = "none"  (similarity)

  i     Eigenvalue     Percent   Cumulative
  1     2.14619361     26.1177     26.1177
  2     1.19032787     14.4830     40.6007
  3     0.84917785     10.3339     50.9346
  4     0.80166166      9.7557     60.6903
  5     0.63906551      7.7770     68.4673
  6     0.45631667      5.5531     74.0204
  7     0.42764289      5.2041     79.2245
  8     0.30567802      3.7199     82.9444
  9     0.29594768      3.6015     86.5459
 10     0.22964513      2.7946     89.3405
 11     0.18839123      2.2926     91.6331
 12     0.17365020      2.1132     93.7463
 13     0.15235256      1.8540     95.6003
 14     0.11426895      1.3906     96.9909
 15     0.07880312      0.9590     97.9499
 16     0.05072685      0.6173     98.5672
 17     0.04345356      0.5288     99.0960
 18     0.03364625      0.4095     99.5054
 19     0.02181415      0.2655     99.7709
 20     0.01324490      0.1612     99.9321
 21     0.00420448      0.0521     99.9842
 22     0.00129795      0.0158    100.0000
 23     0.00000050      0.0000    100.0000
Sum of eigenvalues =       8.217391
Eigenvalues (4 by 4) saved in file: C:\Users\mahmoud\Desktop\bb.NTS
Average root:       0.3573
```

شکل۲۶

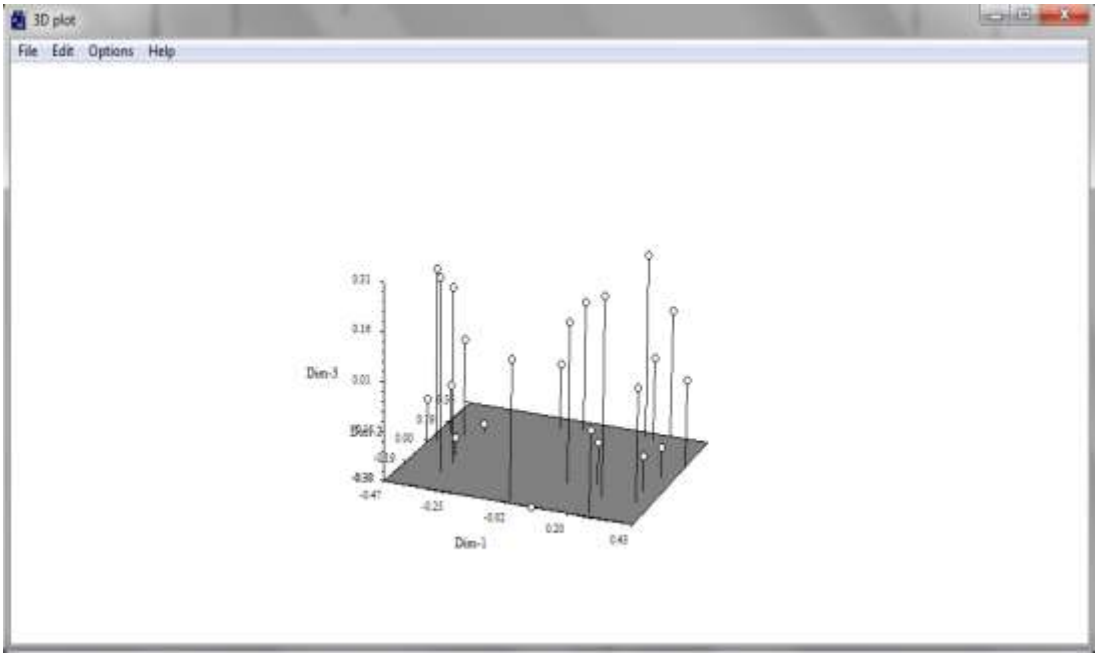
```
Proportions of variance expected using broken-stick model
  Eigenvalue    Expected
           1     16.2361
           2     11.8882
           3      9.7143
           4      8.2650
           5      7.1781
           6      6.3085
           7      5.5839
           8      4.9628
           9      4.4193
          10      3.9362
          11      3.5014
          12      3.1061
          13      2.7438
          14      2.4094
          15      2.0988
          16      1.8090
          17      1.5372
          18      1.2815
          19      1.0399
          20      0.8111
          21      0.5937
          22      0.3867
          23      0.1890
```

شکل۲۷

همچنین دو تا نمودار دوبعدی (شکل۲۸) و سه‌بعدی (شکل۲۹) هم به دست می‌آید.


Matrix plot
File Edit Options Help
Dim-2
Dim-1

شکل۲۸


3D plot
File Edit Options Help
Dim-3
Dim-1

شکل۲۹