

## 4.7 Heatmap 热图

热图 (heatmap) 是“鸢尾花书”中极为常见的可视化手段。特别是在展示数据、矩阵分解时，我们常用热图可视化矩阵。

虽然，matplotlib 中也有绘制热图的工具；但是，推荐大家使用 seaborn 中的 heatmap 函数。这个函数绘制热图更方便。

### Seaborn 中的热图

Seaborn 是一款基于 matplotlib 的数据可视化库，其中包括了各种绘图函数，其中之一就是 heatmap。使用 Seaborn 的 heatmap 函数可以让大家快速而方便地可视化矩阵数据，使得数据分析更加直观和易于理解。

heatmap 可以用于可视化二维数组。图 1 所示为用热图可视化鸢尾花四个量化特征数据。在 Jupyter notebook 中，大家可以看到我们用 cmap 控制色谱，用 xticklabels、yticklabels 分别控制横轴、纵轴标签，用 cbar\_kws 设置色谱条位置，并用 vmin、vmax 控制色谱条起止位置。

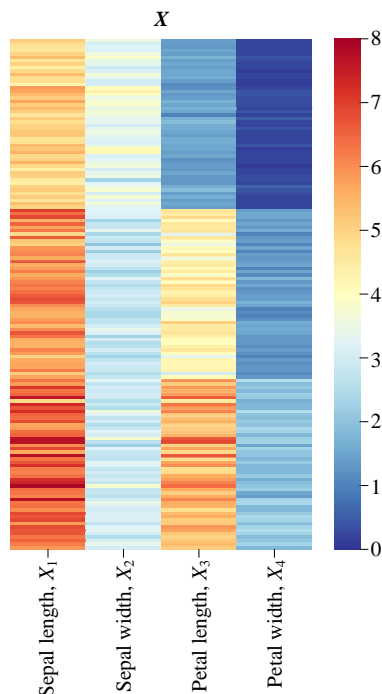


图 1. 热图可视化鸢尾花数据

Seaborn 中的 heatmap 函数还包括许多其他参数，用于自定义热图的外观和行为。例

如，大家可以使用 annot 参数在热图中显示数值，使用 fmt 参数指定数字格式，使用 linewidths 参数调整单元格边框宽度等等。

### 聚类热图

Seaborn 中，clustermap 是一个用于绘制聚类热图的函数，其原理是将矩阵中的行和列进行聚类，并以聚类后的顺序重新排列矩阵的行和列。这样可以将具有相似特征的行和列放在一起，从而更容易地发现它们之间的相似性和差异性。图 2 所示为鸢尾花数据的聚类热图。

“鸢尾花书”的《机器学习》将专门讲解各种聚类算法。

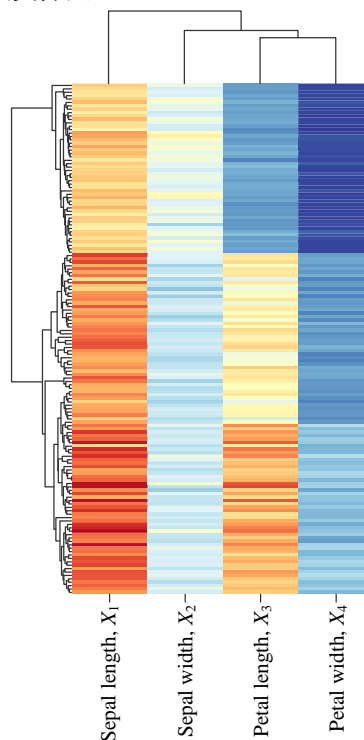


图 2. 热图可视化鸢尾花数据

### 矩阵分解

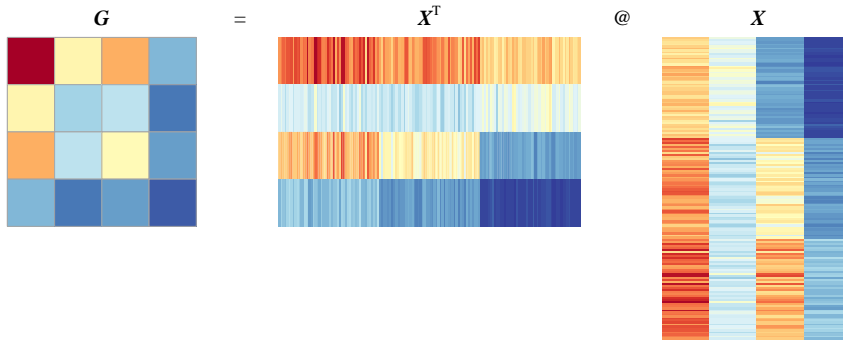
“鸢尾花书”中，大家会经常看到用一组热图可视化矩阵运算，特别是矩阵分解。图 3 所示为常见的几个矩阵运算。注意，后期制作时，热图的形状做了修改。

《矩阵力量》一册将从代数、数据、线性组合、优化、几何、统计等角度和大家讨论这些矩阵运算。

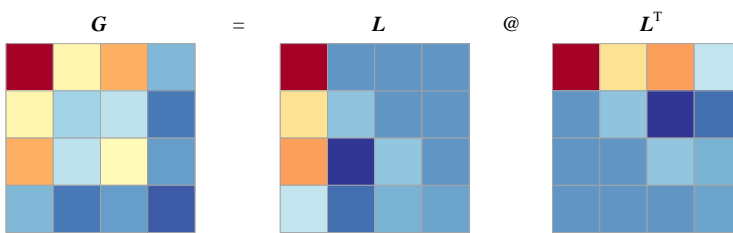
此外，大家还会看到我们用热图可视化协方差矩阵、相关性系数矩阵，以及这些矩阵对应的线性代数运算。本节就不再展开讨论了。

Matplotlib 绘制热图的范例：

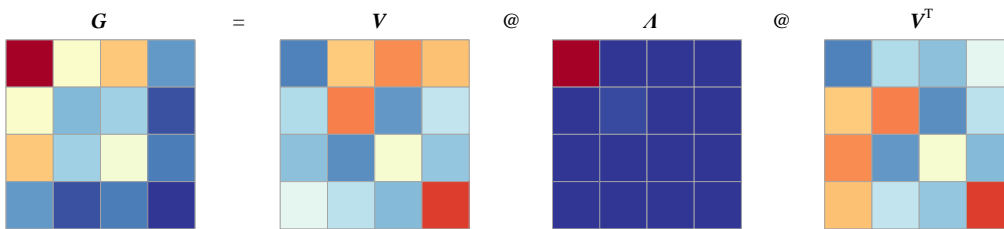
(a) Gram Matrix



(b) Cholesky Decomposition



(c) Eigen Value Decomposition, EVD



(d) Singular Value Decomposition, SVD

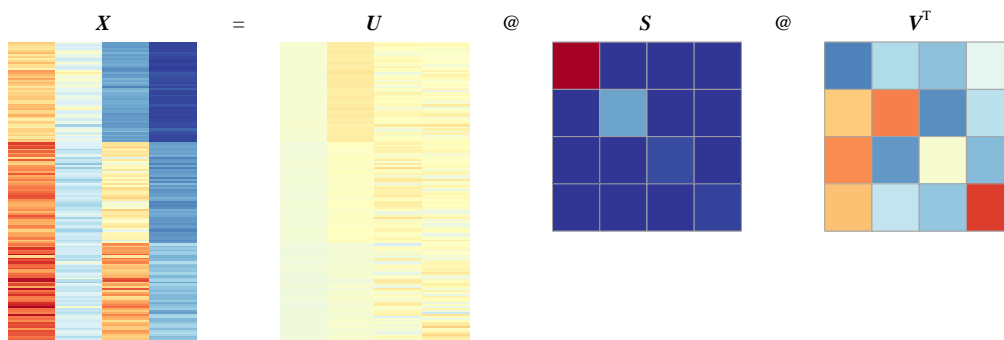


图 3. 用热图可视化矩阵运算

[https://matplotlib.org/stable/gallery/images\\_contours\\_and\\_fields/image\\_annotated\\_heatmap.html](https://matplotlib.org/stable/gallery/images_contours_and_fields/image_annotated_heatmap.html)  
[https://matplotlib.org/stable/gallery/images\\_contours\\_and\\_fields/matshow.html](https://matplotlib.org/stable/gallery/images_contours_and_fields/matshow.html)

多子图如何放置色谱条：

[https://matplotlib.org/stable/gallery/subplots\\_axes\\_and\\_figures/colorbar\\_placement.html](https://matplotlib.org/stable/gallery/subplots_axes_and_figures/colorbar_placement.html)