```
<!DOCTYPE html><html><head>
      <title>lecturenotesul</title>
      <meta charset="utf-8">
      <meta name="viewport" content="width=device-width, initial-
scale=1.0">


        <script type="text/x-mathjax-config">
          MathJax.Hub.Config({"extensions":["tex2jax.js"],"jax":
["input/TeX","output/HTML-CSS"],"messageStyle":"none","tex2jax":
{"processEnvironments":false,"processEscapes":true,"inlineMath":
[["$","$"],["\\(","\\)"]],"displayMath":[["$$","$$"],["\\
[","\\]"]]},"TeX":{"extensions":
["AMSmath.js","AMSsymbols.js","noErrors.js","noUndefined.js"]},"HTML-
CSS":{"availableFonts":["TeX"]}});
        </script>
        <script type="text/javascript" async
src="file:////Users/samuel/.vscode/extensions/shd101wyy.markdown-
preview-enhanced-
0.5.13/node_modules/@shd101wyy/mume/dependencies/mathjax/MathJax.js"
charset="UTF-8"></script>
```

```
      <style>
      /**
 * prism.js Github theme based on GitHub's theme.
 * @author Sam Clarke
 */
code[class*="language-"],
pre[class*="language-"] {
  color: #333;
  background: none;
  font-family: Consolas, "Liberation Mono", Menlo, Courier, monospace;
  text-align: left;
  white-space: pre;
  word-spacing: normal;
  word-break: normal;
  word-wrap: normal;
  line-height: 1.4;

  -moz-tab-size: 8;
  -o-tab-size: 8;
  tab-size: 8;

  -webkit-hyphens: none;
  -moz-hyphens: none;
  -ms-hyphens: none;
  hyphens: none;
```

```
}

/* Code blocks */
pre[class*="language-"] {
  padding: .8em;
  overflow: auto;
  /* border: 1px solid #ddd; */
  border-radius: 3px;
  /* background: #fff; */
  background: #f5f5f5;
}

/* Inline code */
:not(pre) > code[class*="language-"] {
  padding: .1em;
  border-radius: .3em;
  white-space: normal;
  background: #f5f5f5;
}

.token.comment,
.token.blockquote {
  color: #969896;
}

.token.cdata {
  color: #183691;
}

.token.doctype,
.token.punctuation,
.token.variable,
.token.macro.property {
  color: #333;
}

.token.operator,
.token.important,
.token.keyword,
.token.rule,
.token.builtin {
  color: #a71d5d;
}

.token.string,
.token.url,
.token.regex,
.token.attr-value {
  color: #183691;
}

.token.property,
.token.number,
.token.boolean,
.token.entity,
.token.atrule,
.token.constant,
```

```
.token.symbol,
.token.command,
.token.code {
  color: #0086b3;
}

.token.tag,
.token.selector,
.token.prolog {
  color: #63a35c;
}

.token.function,
.token.namespace,
.token.pseudo-element,
.token.class,
.token.class-name,
.token.pseudo-class,
.token.id,
.token.url-reference .token.variable,
.token.attr-name {
  color: #795da3;
}

.token.entity {
  cursor: help;
}

.token.title,
.token.title .token.punctuation {
  font-weight: bold;
  color: #1d3e81;
}

.token.list {
  color: #ed6a43;
}

.token.inserted {
  background-color: #eaffea;
  color: #55a532;
}

.token.deleted {
  background-color: #ffecec;
  color: #bd2c00;
}

.token.bold {
  font-weight: bold;
}

.token.italic {
  font-style: italic;
}
```

```
/* JSON */
.language-json .token.property {
  color: #183691;
}

.language-markup .token.tag .token.punctuation {
  color: #333;
}

/* CSS */
code.language-css,
.language-css .token.function {
  color: #0086b3;
}

/* YAML */
.language-yaml .token.atrule {
  color: #63a35c;
}

code.language-yaml {
  color: #183691;
}

/* Ruby */
.language-ruby .token.function {
  color: #333;
}

/* Markdown */
.language-markdown .token.url {
  color: #795da3;
}

/* Makefile */
.language-makefile .token.symbol {
  color: #795da3;
}

.language-makefile .token.variable {
  color: #183691;
}

.language-makefile .token.builtin {
  color: #0086b3;
}

/* Bash */
.language-bash .token.keyword {
  color: #0086b3;
}

/* highlight */
pre[data-line] {
  position: relative;
  padding: 1em 0 1em 3em;
}
```

```
pre[data-line] .line-highlight-wrapper {
  position: absolute;
  top: 0;
  left: 0;
  background-color: transparent;
  display: block;
  width: 100%;
}

pre[data-line] .line-highlight {
  position: absolute;
  left: 0;
  right: 0;
  padding: inherit 0;
  margin-top: 1em;
  background: hsla(24, 20%, 50%,.08);
  background: linear-gradient(to right, hsla(24, 20%, 50%,.1) 70%,
hsla(24, 20%, 50%,0));
  pointer-events: none;
  line-height: inherit;
  white-space: pre;
}

pre[data-line] .line-highlight:before,
pre[data-line] .line-highlight[data-end]:after {
  content: attr(data-start);
  position: absolute;
  top: .4em;
  left: .6em;
  min-width: 1em;
  padding: 0 .5em;
  background-color: hsla(24, 20%, 50%,.4);
  color: hsl(24, 20%, 95%);
  font: bold 65%/1.5 sans-serif;
  text-align: center;
  vertical-align: .3em;
  border-radius: 999px;
  text-shadow: none;
  box-shadow: 0 1px white;
}

pre[data-line] .line-highlight[data-end]:after {
  content: attr(data-end);
  top: auto;
  bottom: .4em;
}html body{font-family:"Helvetica Neue",Helvetica,"Segoe
UI",Arial,freesans,sans-serif;font-size:16px;line-
height:1.6;color:#333;background-color:#fff;overflow:initial;box-
sizing:border-box;word-wrap:break-word}html body>:first-child{margin-
top:0}html body h1,html body h2,html body h3,html body h4,html body
h5,html body h6{line-height:1.2;margin-top:1em;margin-
bottom:16px;color:#000}html body h1{font-size:2.25em;font-
weight:300;padding-bottom:.3em}html body h2{font-size:1.75em;font-
weight:400;padding-bottom:.3em}html body h3{font-size:1.5em;font-
weight:500}html body h4{font-size:1.25em;font-weight:600}html body
h5{font-size:1.1em;font-weight:600}html body h6{font-size:1em;font-
weight:600}html body h1,html body h2,html body h3,html body h4,html
```

body h5{font-weight:600}html body h5{font-size:1em}html body h6{color:#5c5c5c}html body strong{color:#000}html body del{color:#5c5c5c}html body a:not([href]){color:inherit;text-decoration:none}html body a{color:#08c;text-decoration:none}html body a:hover{color:#00a3f5;text-decoration:none}html body img{max-width:100%}html body>p{margin-top:0;margin-bottom:16px;word-wrap:break-word}html body>ul,html body>ol{margin-bottom:16px}html body ul,html body ol{padding-left:2em}html body ul.no-list,html body ol.no-list{padding:0;list-style-type:none}html body ul ul,html body ul ol,html body ol ol,html body ol ul{margin-top:0;margin-bottom:0}html body li{margin-bottom:0}html body li.task-list-item{list-style:none}html body li>p{margin-top:0;margin-bottom:0}html body .task-list-item-checkbox{margin:0 .2em .25em -1.8em;vertical-align:middle}html body .task-list-item-checkbox:hover{cursor:pointer}html body blockquote{margin:16px 0;font-size:inherit;padding:0 15px;color:#5c5c5c;background-color:#f0f0f0;border-left:4px solid #d6d6d6}html body blockquote>:first-child{margin-top:0}html body blockquote>:last-child{margin-bottom:0}html body hr{height:4px;margin:32px 0;background-color:#d6d6d6;border:0 none}html body table{margin:10px 0 15px 0;border-collapse:collapse;border-spacing:0;display:block;width:100%;overflow:auto;word-break:normal;word-break:keep-all}html body table th{font-weight:bold;color:#000}html body table td,html body table th{border:1px solid #d6d6d6;padding:6px 13px}html body dl{padding:0}html body dl dt{padding:0;margin-top:16px;font-size:1em;font-style:italic;font-weight:bold}html body dl dd{padding:0 16px;margin-bottom:16px}html body code{font-family:Menlo,Monaco,Consolas,'Courier New',monospace;font-size:.85em !important;color:#000;background-color:#f0f0f0;border-radius:3px;padding:.2em 0}html body code::before,html body code::after{letter-spacing:-0.2em;content:"\00a0"}html body pre>code{padding:0;margin:0;font-size:.85em !important;word-break:normal;white-space:pre;background:transparent;border:0}html body .highlight{margin-bottom:16px}html body .highlight pre,html body pre{padding:1em;overflow:auto;font-size:.85em !important;line-height:1.45;border:#d6d6d6;border-radius:3px}html body .highlight pre{margin-bottom:0;word-break:normal}html body pre code,html body pre tt{display:inline;max-width:initial;padding:0;margin:0;overflow:initial;line-height:inherit;word-wrap:normal;background-color:transparent;border:0}html body pre code:before,html body pre tt:before,html body pre code:after,html body pre tt:after{content:normal}html body p,html body blockquote,html body ul,html body ol,html body dl,html body pre{margin-top:0;margin-bottom:16px}html body kbd{color:#000;border:1px solid #d6d6d6;border-bottom:2px solid #c7c7c7;padding:2px 4px;background-color:#f0f0f0;border-radius:3px}@media print{html body{background-color:#fff}html body h1,html body h2,html body h3,html body h4,html body h5,html body h6{color:#000;page-break-after:avoid}html body blockquote{color:#5c5c5c}html body pre{page-break-inside:avoid}html body table{display:table}html body img{display:block;max-width:100%;max-height:100%}html body pre,html body code{word-wrap:break-word;white-space:pre}}.markdown-preview{width:100%;height:100%;box-sizing:border-box}.markdown-preview .pagebreak,.markdown-preview .newpage{page-break-before:always}.markdown-preview pre.line-

numbers{position:relative;padding-left:3.8em;counter-reset:linenumber}.markdown-preview pre.line-numbers>code{position:relative}.markdown-preview pre.line-numbers .line-numbers-rows{position:absolute;pointer-events:none;top:1em;font-size:100%;left:0;width:3em;letter-spacing:-1px;border-right:1px solid #999;-webkit-user-select:none;-moz-user-select:none;-ms-user-select:none;user-select:none}.markdown-preview pre.line-numbers .line-numbers-rows>span{pointer-events:none;display:block;counter-increment:linenumber}.markdown-preview pre.line-numbers .line-numbers-rows>span:before{content:counter(linenumber);color:#999;display:block;padding-right:.8em;text-align:right}.markdown-preview .mathjax-exps .MathJax_Display{text-align:center !important}.markdown-preview:not([for="preview"]) .code-chunk .btn-group{display:none}.markdown-preview:not([for="preview"]) .code-chunk .status{display:none}.markdown-preview:not([for="preview"]) .code-chunk .output-div{margin-bottom:16px}.scrollbar-style::-webkit-scrollbar{width:8px}.scrollbar-style::-webkit-scrollbar-track{border-radius:10px;background-color:transparent}.scrollbar-style::-webkit-scrollbar-thumb{border-radius:5px;background-color:rgba(150,150,150,0.66);border:4px solid rgba(150,150,150,0.66);background-clip:content-box}html body[for="html-export"]:not([data-presentation-mode]) {position:relative;width:100%;height:100%;top:0;left:0;margin:0;padding:0;overflow:auto}html body[for="html-export"]:not([data-presentation-mode]) .markdown-preview{position:relative;top:0}@media screen and (min-width:914px){html body[for="html-export"]:not([data-presentation-mode]) .markdown-preview{padding:2em calc(50% - 457px + 2em)}}@media screen and (max-width:914px){html body[for="html-export"]:not([data-presentation-mode]) .markdown-preview{padding:2em}}@media screen and (max-width:450px){html body[for="html-export"]:not([data-presentation-mode]) .markdown-preview{font-size:14px !important;padding:1em}}@media print{html body[for="html-export"]:not([data-presentation-mode]) #sidebar-toc-btn{display:none}}html body[for="html-export"]:not([data-presentation-mode]) #sidebar-toc-btn{position:fixed;bottom:8px;left:8px;font-size:28px;cursor:pointer;color:inherit;z-index:99;width:32px;text-align:center;opacity:.4}html body[for="html-export"]:not([data-presentation-mode])[html-show-sidebar-toc] #sidebar-toc-btn{opacity:1}html body[for="html-export"]:not([data-presentation-mode])[html-show-sidebar-toc] .md-sidebar-toc{position:fixed;top:0;left:0;width:300px;height:100%;padding:32px 0 48px 0;font-size:14px;box-shadow:0 0 4px rgba(150,150,150,0.33);box-sizing:border-box;overflow:auto;background-color:inherit}html body[for="html-export"]:not([data-presentation-mode])[html-show-sidebar-toc] .md-sidebar-toc::-webkit-scrollbar{width:8px}html body[for="html-export"]:not([data-presentation-mode])[html-show-sidebar-toc] .md-sidebar-toc::-webkit-scrollbar-track{border-radius:10px;background-color:transparent}html body[for="html-export"]:not([data-presentation-mode])[html-show-sidebar-toc] .md-sidebar-toc::-webkit-scrollbar-thumb{border-radius:5px;background-color:rgba(150,150,150,0.66);border:4px solid rgba(150,150,150,0.66);background-clip:content-box}html body[for="html-export"]:not([data-presentation-mode])[html-show-sidebar-toc] .md-sidebar-toc a{text-decoration:none}html body[for="html-export"]:not([data-presentation-mode])[html-show-sidebar-toc] .md-sidebar-toc ul{padding:0 1.6em;margin-top:.8em}html body[for="html-export"]:not([data-presentation-mode])[html-show-

sidebar-toc] .md-sidebar-toc li{margin-bottom:.8em}html
body[for="html-export"]:not([data-presentation-mode])[html-show-
sidebar-toc] .md-sidebar-toc ul{list-style-type:none}html
body[for="html-export"]:not([data-presentation-mode])[html-show-
sidebar-toc] .markdown-preview{left:300px;width:calc(100% -
300px);padding:2em calc(50% - 457px - 150px);margin:0;box-
sizing:border-box}@media screen and (max-width:1274px){html
body[for="html-export"]:not([data-presentation-mode])[html-show-
sidebar-toc] .markdown-preview{padding:2em}}@media screen and (max-
width:450px){html body[for="html-export"]:not([data-presentation-
mode])[html-show-sidebar-toc] .markdown-preview{width:100%}}html
body[for="html-export"]:not([data-presentation-mode]):not([html-show-
sidebar-toc]) .markdown-
preview{left:50%;transform:translateX(-50%)}html body[for="html-
export"]:not([data-presentation-mode]):not([html-show-sidebar-toc])
.md-sidebar-toc{display:none}
/* Please visit the URL below for more information: */
/*   https://shd101wyy.github.io/markdown-preview-
enhanced/#/customize-css */
.markdown-preview.markdown-preview h1,
.markdown-preview.markdown-preview h2,
.markdown-preview.markdown-preview h3,
.markdown-preview.markdown-preview h4,
.markdown-preview.markdown-preview h5,
.markdown-preview.markdown-preview h6 {
  font-weight: bolder;
  text-decoration-line: underline;
}

```
      </style>
    </head>
    <body for="html-export">
      <div class="mume markdown-preview  ">
      <h1 class="mume-header" id="unsupervised-learning">Unsupervised
Learning</h1>

<h2 class="mume-header" id="examples-of-unsupervised-
learning">Examples of Unsupervised Learning</h2>

<ul>
<li>Compared to Supervised Learning, Unsupervised learning deals with
data with no specific target variable / ground truth
<ul>
<li>The lack of &amp;quot;ground truth&amp;quot; means we don&amp;apos;t have a
measurement of &amp;quot;accuracy&amp;quot;, or precision / recall, or R-
squared</li>
</ul>
</li>
<li>Market Basket Analysis</li>
<li>Spotify Music Recommendation</li>
<li>Customer Segmentation (Clustering)</li>
<li>Anomaly Transactions (Anomaly Detection)
<ul>
<li>We don&amp;apos;t know if a transaction is from a stolen credit card
or has fraudulent intents until after the fact (post-investigation)
</li>
</ul>
```


Text auto-completion
Dimensionality Reduction

Today we live in an era of "big data", almost everything around us are getting larger in dimensionality
Curse of dimensionality

<h2 class="mume-header" id="dimensionality-reduction">Dimensionality Reduction</h2>

To reduce dimensions, we can do one of two things:

Feature Selection

Eliminate any features that are not highly predictive, elimiate highly correlated variables
Examples: Stepwise backward elimination, remove near zero variance

Principal Components

Doesn't eliminate any features; It merely summarizes the features in new dimensions (this seeks to retain as much information as possible using as little dimensions as possible)

In the case of dimensionality reduction, variance is our information (the more variance we can pack into our principal components, the better); Correlation is redundancy (the less correlation, the better)

If Body fat percentage can accurately predict BMI (body mass index) because both are highly correlated, then storing both variables are "redundant". We can use one to predict the other

Fundamentals of linear algebra: scalar, vector and matrix. A scalar when acted on the vector, only scales the vector, does not change its direction. A matrix, as a rule, when acted on a vector, changes the direction and the scale of the vector.

Exceptions: Identity matrix, rotation matrix, and another exception
As a rule, for any given matrix there exists certain vectors where the matrix can act on said vectors and do not change the vector's direction but only scales it

We give these vectors names, we call them
"eigenvectors", and the scalar "eigenvalues"

<pre data-role="codeBlock" data-info="r" class="language-r">A <span
class="token operator"><- matrix<span class="token
punctuation">(...<span
class="token punctuation">)
eigen(A<span class="token
punctuation">)$vectors
eigen(A<span class="token
punctuation">)$values
</pre>

<h3 class="mume-header" id="principal-components-in-
practice">Principal Components in Practice</h3>

<pre data-role="codeBlock" data-info="r" class="language-r"><span
class="token comment"># are my variables in the same scale
model1 <- prcomp<span
class="token punctuation">(data<span class="token
punctuation">, scale=<span
class="token boolean">TRUE)

plot(model1<span class="token
punctuation">)
biplot(model1<span class="token
punctuation">)
learn to interpret the eigenvalues (%
variance explained)
summary(model1<span
class="token punctuation">)
</pre> You can also use the <code>FactoMineR</code> package:
<pre data-role="codeBlock" data-info="r" class="language-
r">library(<span class="token
string">"FactoMineR")

model2 <- PCA<span class="token
punctuation">(data,
quali.sup=c<span class="token
punctuation">(2<span
class="token punctuation">,4
,<span class="token
number">5)<span
class="token punctuation">, graph=
FALSE<span class="token
punctuation">)
plot(model2<span class="token
punctuation">, choix=<span
class="token string">"ind"<span class="token
punctuation">, habillage=
3,
 select=c<span class="token
punctuation">("Okinawa"
, <span class="token
string">"Cesario")

```
</span><span class="token punctuation">)</span>
plot<span class="token punctuation">(</span>model2<span class="token
punctuation">,</span> choix<span class="token operator">=</span><span
class="token string">&quot;ind&quot;</span><span class="token
punctuation">,</span> habillage<span class="token operator">=</span>
<span class="token number">3</span><span class="token punctuation">,
</span> select<span class="token operator">=</span><span class="token
number">1</span><span class="token operator">:</span><span
class="token number">10</span><span class="token punctuation">)</span>
plot<span class="token punctuation">(</span>model2<span class="token
punctuation">,</span> choix<span class="token operator">=</span><span
class="token string">&quot;ind&quot;</span><span class="token
punctuation">,</span> habillage<span class="token operator">=</span>
<span class="token number">3</span><span class="token punctuation">,
</span> select<span class="token operator">=</span><span class="token
string">&quot;contrib10&quot;</span><span class="token punctuation">)
</span>
plot<span class="token punctuation">(</span>model2<span class="token
punctuation">,</span> choix<span class="token operator">=</span><span
class="token string">&quot;var&quot;</span><span class="token
punctuation">)</span>

reconstructed <span class="token operator">&lt;-</span> reconst<span
class="token punctuation">(</span>model2<span class="token
punctuation">,</span> ncp<span class="token operator">=</span><span
class="token number">4</span><span class="token punctuation">)</span>
</pre><p>Principal components have a lot of other uses, among them:
</p>
<ul>
<li>Anomaly Detection</li>
<li>Embedding Projector by TensorFlow</li>
<li>How it&apos;s used to visualize language models that are otherwise
high-dimensional (elang example)</li>
</ul>

      </div>
      <div class="md-sidebar-toc"><ul>
<li><a href="#unsupervised-learning">Unsupervised Learning</a>
<ul>
<li><a href="#examples-of-unsupervised-learning">Examples of
Unsupervised Learning</a></li>
<li><a href="#dimensionality-reduction">Dimensionality Reduction</a>
<ul>
<li><a href="#principal-components-in-practice">Principal Components
in Practice</a></li>
</ul>
</li>
</ul>
</li>
</ul>
</div>
      <a id="sidebar-toc-btn">&#x2261;</a>
```

```
<script>

var sidebarTOCBtn = document.getElementById('sidebar-toc-btn')
sidebarTOCBtn.addEventListener('click', function(event) {
  event.stopPropagation()
  if (document.body.hasAttribute('html-show-sidebar-toc')) {
    document.body.removeAttribute('html-show-sidebar-toc')
  } else {
    document.body.setAttribute('html-show-sidebar-toc', true)
  }
})
</script>


    </body></html>
```