COMPUTATIONAL FINANCE & RISK MANAGEMENT
UNIVERSITY *of* WASHINGTON
Department of Applied Mathematics

AMATH 460: Mathematical Methods for Quantitative Finance

6. Linear Algebra II

Kjell Konis
Acting Assistant Professor, Applied Mathematics
University of Washington

Outline

Outline

Transposes

- Let A be an $m \times n$ matrix, the transpose of A is denoted by $A^\intercal$
- The columns of $A^\intercal$ are the rows of A
- If the dimension of A is $m \times n$ then the dimension of $A^\intercal$ is $n \times m$

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \qquad A^\intercal = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Properties

- $(A^\intercal)^\intercal = A$
- The transpose of $A + B$ is $A^\intercal + B^\intercal$
- The transpose of AB is $(AB)^\intercal = B^\intercal A^\intercal$
- The transpose of A^{-1} is $(A^{-1})^\intercal = (A^\intercal)^{-1}$

Symmetric Matrices

- A symmetric matrix satisfies $A = A^\mathsf{T}$
- Implies that $a_{ij} = a_{ji}$

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 4 \\ 3 & 4 & 9 \end{bmatrix} = A^\mathsf{T}$$

- A diagonal matrix is automatically symmetric

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 9 \end{bmatrix} = D^\mathsf{T}$$

Products $R^\top R$, $RR^\top$ and $LDL^\top$

- Let R be any $m \times n$ matrix
- The matrix $A = R^\top R$ is a symmetric matrix

$$A^\top = (R^\top R)^\top = R^\top (R^\top)^\top = R^\top R = A$$

- The matrix $A = RR^\top$ is also a symmetric matrix

$$A^\top = (RR^\top)^\top = (R^\top)^\top R^\top = RR^\top = A$$

- Many problems that start with a rectangular matrix R end up with $R^\top R$ or $RR^\top$ or both!

Permutation Matrices

- An $n \times n$ permutation matrix P has the rows of I in any order
- There are 6 possible 3×3 permutation matrices

$$I = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \quad P_{21} = \begin{bmatrix} & 1 & \\ 1 & & \\ & & 1 \end{bmatrix} \quad P_{32}P_{21} = \begin{bmatrix} & 1 & \\ & & 1 \\ 1 & & \end{bmatrix}$$

$$P_{31} = \begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix} \quad P_{32} = \begin{bmatrix} 1 & & \\ & & 1 \\ & 1 & \end{bmatrix} \quad P_{21}P_{32} = \begin{bmatrix} & & 1 \\ 1 & & \\ & 1 & \end{bmatrix}$$

- P^{-1} is the same as P^{T}
- Example

$$P_{32} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \qquad P_{32} \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Outline

Spaces of Vectors

- The space $\mathbb{R}^n$ consists of all column vectors with n components
- For example, $\mathbb{R}^3$ all column vectors with 3 components and is called 3-dimensional space
- The space $\mathbb{R}^2$ is the xy plane: the two components of the vector are the x and y coordinates and the tail starts at the origin $(0, 0)$
- Two essential vector operations go on inside the vector space:
  - Add two vectors in $\mathbb{R}^n$
  - Multiply a vector in $\mathbb{R}^n$ by a scalar
- The result lies in the same vector space $\mathbb{R}^n$

Subspaces

- A subspace of a vector space is a set of vectors (including the zero vector) that satisfies two requirements:

  If u and w are vectors in the subspace and c is any scalar, then

  (i) $u + w$ is in the subspace

  (ii) cw is in the subspace

- Some subspaces of $\mathbb{R}^3$

  L Any line through $(0, 0, 0)$, e.g., the x axis

  P Any plane through $(0, 0, 0)$, e.g., the xy plane

  $\mathbb{R}^3$ The whole space

  Z The zero vector

The Column Space of A

- The column space of a matrix A consists of all linear combinations of its columns
- The linear combinations are vectors that can be written as Ax
- $Ax = b$ is solvable if and only if b is in the column space of A
- Let A be an $m \times n$ matrix
  - The columns of A have m components
  - The columns of A live in $\mathbb{R}^m$
  - The column space of A is a subspace of $\mathbb{R}^m$
- The column space of A is denoted by $R(A)$
  - R stands for range

The Nullspace of A

- The <u>nullspace</u> of A consists of <u>all solutions</u> to $Ax = 0$
- The nullspace of A is denoted by $N(A)$
- Example: elimination

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 8 \\ 3 & 6 & 11 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

  - The <u>pivot</u> variables are x_1 and x_3
  - The <u>free</u> variable is x_2 (column 2 has no pivot)

- The number of pivots is called the <u>rank</u> of A

Linear Independence

- A sequence of vectors $v_1, v_2, \ldots, v_n$ is linearly independent if the only linear combination that gives the zero vector is $0v_1 + 0v_2 + \ldots + 0v_n$
- The columns of A are independent if the only solution to $Ax = 0$ is the zero vector
  - Elimination produces no free variables
  - The rank of A is equal to n
  - The nullspace $N(A)$ contains only the zero vector
- A set of vectors <u>spans</u> a space if their linear combinations fill the space
- A <u>basis</u> for a vector space is a sequence of vectors that
  (i) are linearly independent
  (ii) span the space
- The <u>dimension of a vector space</u> is the number of vectors in every basis

Outline

Sample Variance and Covariance

- <u>Sample mean</u> of the elements of a vector x of length m

$$\bar{x} = \frac{1}{m} \sum_{i=1}^{m} x_i$$

- <u>Sample variance</u> of the elements of x

$$\mathrm{Var}(x) = \frac{1}{m-1} \sum_{i=1}^{m} (x_i - \bar{x})^2$$

- Let y be vector of length m
- <u>Sample covariance</u> of x and y

$$\mathrm{Cov}(x, y) = \frac{1}{m-1} \sum_{i=1}^{m} (x_i - \bar{x})(y_i - \bar{y})$$

Sample Variance

- Let e be a column vector of m ones

$$\frac{e^{\mathsf{T}} x}{m} = \frac{1}{m} \sum_{i=1}^{m} 1 x_i = \frac{1}{m} \sum_{i=1}^{m} x_i = \bar{x}$$

- $\bar{x}$ is a scalar
- Let $e\,\bar{x}$ be a column vector repeating $\bar{x}$ m times
- Let $\tilde{x} = x - e\frac{e^{\mathsf{T}} x}{m} = x - e\,\bar{x}$
- The i^{th} element of $\tilde{x}$ is

$$\tilde{x}_i = x_i - \bar{x}$$

- Take another look at the sample variance

$$\text{Var}(x) = \frac{1}{m-1} \sum_{i=1}^{m} (x_i - \bar{x})^2 = \frac{1}{m-1} \sum_{i=1}^{m} \tilde{x}_i^2 = \frac{\tilde{x}^{\mathsf{T}} \tilde{x}}{m-1} = \frac{\|\tilde{x}\|^2}{m-1}$$

Sample Covariance

- A similar result holds for the sample covariance
- Let $\tilde{y} = y - e\dfrac{e^{\mathsf{T}}y}{m} = y - e\,\bar{y}$
- The sample covariance becomes

$$\mathrm{Cov}(x, y) = \frac{1}{m-1}\sum_{i=1}^{m}(x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{m-1}\sum_{i=1}^{m}\tilde{x}_i\tilde{y}_i \;=\; \frac{\tilde{x}^{\mathsf{T}}\tilde{y}}{m-1}$$

- Observe that $\mathrm{Var}(x) = \mathrm{Cov}(x, x)$
- Proceed with $\mathrm{Cov}(x, x)$ and treat $\mathrm{Var}(x)$ as a special case

Variance-Covariance Matrix

- Suppose x and y are the columns of a matrix R

$$R = \begin{bmatrix} | & | \\ x & y \\ | & | \end{bmatrix} \qquad \tilde{R} = \begin{bmatrix} | & | \\ \tilde{x} & \tilde{y} \\ | & | \end{bmatrix}$$

- The sample variance-covariance matrix is

$$\mathrm{Cov}(R) = \begin{bmatrix} \mathrm{Cov}(x,x) & \mathrm{Cov}(x,y) \\ \mathrm{Cov}(y,x) & \mathrm{Cov}(y,y) \end{bmatrix} = \frac{1}{m-1} \begin{bmatrix} \tilde{x}^\mathsf{T}\tilde{x} & \tilde{x}^\mathsf{T}\tilde{y} \\ \tilde{y}^\mathsf{T}\tilde{x} & \tilde{y}^\mathsf{T}\tilde{y} \end{bmatrix} = \frac{\tilde{R}^\mathsf{T}\tilde{R}}{m-1}$$

- The sample variance-covariance matrix is symmetric

Outline

Computing Covariance Matrices

- Take a closer look at how $\tilde{x}$ was computed

$$\tilde{x} = x - e\,\bar{x} = x - e\frac{e^{\mathsf{T}}x}{m} = x - \frac{ee^{\mathsf{T}}}{m}x = \underbrace{\left(I - \frac{ee^{\mathsf{T}}}{m}\right)}_{A}x$$

- What is A?
  - The outer product (ee^{T}) is an $m \times m$ matrix
  - I is the $m \times m$ identity matrix
  - A is an $m \times m$ matrix
- Premultiplication by A turns x into $\tilde{x}$
- Can think of matrix multiplication as one matrix acting on another

Computing Covariance Matrices

- Next, consider what happens when we premultiply R by A
- Think of R in block structure where each column is a block

$$AR = A \begin{bmatrix} | & | \\ x & y \\ | & | \end{bmatrix} = \begin{bmatrix} | & | \\ Ax & Ay \\ | & | \end{bmatrix} = \begin{bmatrix} | & | \\ \tilde{x} & \tilde{y} \\ | & | \end{bmatrix} = \tilde{R}$$

- Expression for the variance-covariance matrix no longer needs $\tilde{R}$

$$\mathrm{Cov}(R) = \frac{1}{n-1}\tilde{R}^{\mathsf{T}}\tilde{R} = \frac{1}{m-1}(AR)^{\mathsf{T}}(AR)$$

- Since R has 2 columns, $\mathrm{Cov}(R)$ is a 2×2 matrix
- In general, R may have n columns $\Longrightarrow$ $\mathrm{Cov}(R)$ is an $n \times n$ matrix
  - Still use the same $m \times m$ matrix A

Computing Covariance Matrices

- Take another look at formula for variance-covariance matrix
- Rule for transpose of a product

$$\text{Cov}(R) = \frac{1}{m-1}(AR)^{\mathsf{T}}(AR) = \frac{1}{m-1}R^{\mathsf{T}}A^{\mathsf{T}}AR$$

- Consider the product

$$\begin{aligned}
A^{\mathsf{T}}A &= \left(I - \frac{ee^{\mathsf{T}}}{m}\right)^{\mathsf{T}}\left(I - \frac{ee^{\mathsf{T}}}{m}\right) \\
&= \left[I^{\mathsf{T}} - \left(\frac{ee^{\mathsf{T}}}{m}\right)^{\mathsf{T}}\right]\left(I - \frac{ee^{\mathsf{T}}}{m}\right) \\
&= \left[I^{\mathsf{T}} - \frac{[e^{\mathsf{T}}]^{\mathsf{T}}e^{\mathsf{T}}}{m}\right]\left(I - \frac{ee^{\mathsf{T}}}{m}\right) \\
&= \left(I - \frac{ee^{\mathsf{T}}}{m}\right)\left(I - \frac{ee^{\mathsf{T}}}{m}\right) \qquad \text{Since } A^{\mathsf{T}} = A,\ A \text{ is symmetric}
\end{aligned}$$

Computing Covariance Matrices

- Continuing ...

$$
\begin{aligned}
A^{\mathsf{T}}A &= \left(I - \frac{ee^{\mathsf{T}}}{m}\right)\left(I - \frac{ee^{\mathsf{T}}}{m}\right) = \left(I - \frac{ee^{\mathsf{T}}}{m}\right)^2 = A^2 \\
&= I^2 - \frac{ee^{\mathsf{T}}}{m} - \frac{ee^{\mathsf{T}}}{m} + \left(\frac{ee^{\mathsf{T}}}{m}\right)\left(\frac{ee^{\mathsf{T}}}{m}\right) \\
&= I - 2\frac{ee^{\mathsf{T}}}{m} + \frac{e(e^{\mathsf{T}}e)e^{\mathsf{T}}}{m^2} \\
&= I - 2\frac{ee^{\mathsf{T}}}{m} + \frac{eme^{\mathsf{T}}}{m^2} \\
&= \left(I - \frac{ee^{\mathsf{T}}}{m}\right) = A
\end{aligned}
$$

- A matrix satisfying $A^2 = A$ is called <u>idempotent</u>

Computing Covariance Matrices

- Can simplify the expression for the sample variance-covariance matrix

$$\mathrm{Cov}(R) = \frac{1}{m-1} R^{\mathsf{T}} \left[A^{\mathsf{T}} A\right] R = \underbrace{\frac{1}{m-1}}_{\text{scalar}} \; \underbrace{R^{\mathsf{T}}}_{n\times m} \; \underbrace{A}_{m\times m} \; \underbrace{R}_{m\times n}$$

- How to order the operations ...

$$\begin{aligned} R^{\mathsf{T}} A R \quad &= \quad R^{\mathsf{T}} \left(I - \frac{e e^{\mathsf{T}}}{m}\right) R = R^{\mathsf{T}} \left(R - \frac{1}{m} e \overbrace{\underbrace{e^{\mathsf{T}}}_{1\times m} \underbrace{R}_{m\times n}}^{1\times n}\right) \\ &= \quad R^{\mathsf{T}} \left(R - \frac{1}{m} \underbrace{e}_{m\times 1} \underbrace{M_1}_{1\times n}\right) = \underbrace{R^{\mathsf{T}}}_{n\times m} \overbrace{\left(\underbrace{R}_{m\times n} - \frac{1}{m} \underbrace{M_2}_{m\times n}\right)}^{m\times n} = \underbrace{\mathrm{Cov}(R)}_{n\times n} \end{aligned}$$

Outline

Orthogonal Matrices

- Two vectors $q, w \in \mathbb{R}^m$ are orthogonal if their inner product is zero

$$\langle q, w \rangle = 0$$

- Consider a set of m vectors $\{q_1, \ldots, q_m\}$ where $q_j \in \mathbb{R}^m \setminus 0$
- Assume that the vectors $\{q_1, \ldots, q_m\}$ are pairwise orthogonal
- Let $\tilde{q}_j = \dfrac{q_j}{\|q_j\|}$ be a unit vector in the same direction as q_j
- The vectors $\{\tilde{q}_a, \ldots, \tilde{q}_m\}$ are <u>orthonormal</u>
- Let Q be a matrix with columns $\{\tilde{q}_j\}$, consider the product QQ^T

$$Q^\mathsf{T} Q = \begin{bmatrix} — & \tilde{q}_1^\mathsf{T} & — \\ \cdots & \cdots & \cdots \\ — & \tilde{q}_m^\mathsf{T} & — \end{bmatrix} \begin{bmatrix} | & \vdots & | \\ \tilde{q}_1 & \vdots & \tilde{q}_m \\ | & \vdots & | \end{bmatrix} = \begin{bmatrix} \tilde{q}_1^\mathsf{T} \tilde{q}_1 & & \tilde{q}_i^\mathsf{T} \tilde{q}_j \\ & \ddots & \\ \tilde{q}_i^\mathsf{T} \tilde{q}_j & & \tilde{q}_m^\mathsf{T} \tilde{q}_m \end{bmatrix} = I$$

Orthogonal Matrices

- A square matrix Q is <u>orthogonal</u> if $Q^{\mathsf{T}} Q = I$ and $Q Q^{\mathsf{T}} = I$
- Orthogonal matrices represent rotations and reflections
- Example:

$$Q = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

Rotates a vector in the xy plane through the angle θ

$$Q^{\mathsf{T}} Q = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

$$= \begin{bmatrix} \cos^2(\theta) + \sin^2(\theta) & -\cos(\theta)\sin(\theta) + \sin(\theta)\cos(\theta) \\ -\sin(\theta)\cos(\theta) + \cos(\theta)\sin(\theta) & \sin^2(\theta) + \cos^2(\theta) \end{bmatrix} = I$$

Properties of Orthogonal Matrices

- The definition $Q^{\mathsf{T}}Q = QQ^{\mathsf{T}} = I$ implies $Q^{-1} = Q^{\mathsf{T}}$

$$Q^{\mathsf{T}} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix}$$

Cosine is an even function and sine is an odd function

- Multiplication by an orthogonal matrix Q preserves dot products

$$(Qx) \cdot (Qy) = (Qx)^{\mathsf{T}}(Qy) = x^{\mathsf{T}}Q^{\mathsf{T}}Qy = x^{\mathsf{T}}Iy = x^{\mathsf{T}}y = x \cdot y$$

- Multiplication by an orthogonal matrix Q leaves lengths unchanged

$$\|Qx\| = \|x\|$$

Outline

Singular Value Factorization

- So far ...
  - If Q is orthogonal then $Q^{-1} = Q^{\mathsf{T}}$
  - if D is diagonal then D^{-1} is

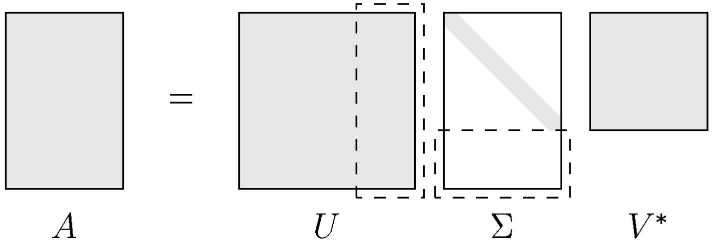
$$D = \begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_m \end{bmatrix} \qquad D^{-1} = \begin{bmatrix} 1/d_1 & & 0 \\ & \ddots & \\ 0 & & 1/d_m \end{bmatrix}$$

  - Orthogonal and diagonal matrices have nice properties

- Wouldn't it be nice if any matrix could be expressed as a product of diagonal and orthogonal matrices ...

Singular Value Factorization

- Every $m \times n$ matrix A can be factored into $A = U\Sigma V^\intercal$ where

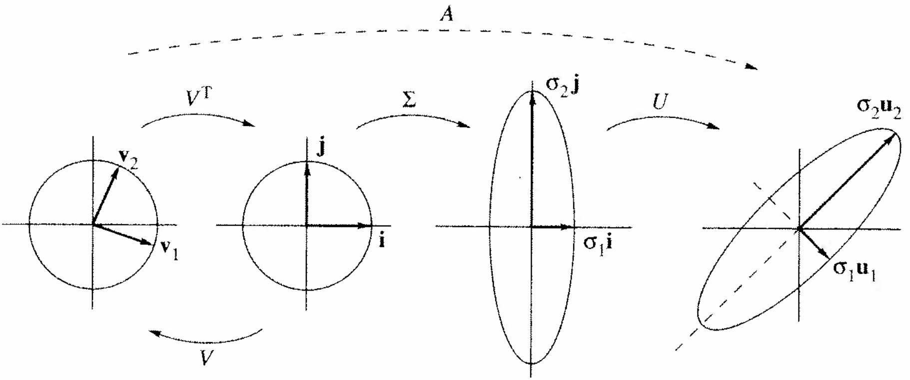


Picture for $m > n$

- U is an $m \times m$ orthogonal matrix whose columns are the left singular vectors of A
- Σ is an $m \times n$ diagonal matrix containing the singular values of A
  - Convention: $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$
- V is an $n \times n$ orthogonal matrix whose columns contain the right singular vectors of A

Multiplication

- Every invertible 2×2 matrix transforms the unit circle into an ellipse

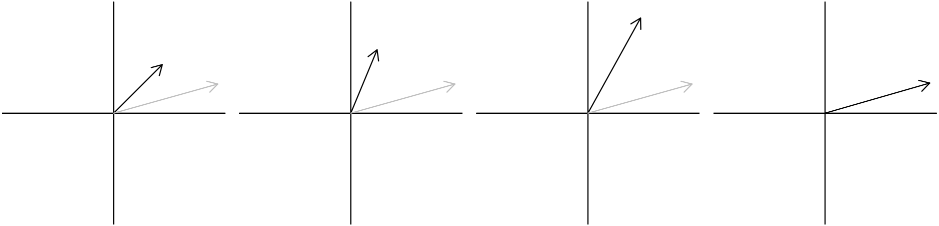


$$Av_2 = U\Sigma V^\top v_2 = U\Sigma \begin{bmatrix} -v_1^\top- \\ -v_2^\top- \end{bmatrix} v_2 = U \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} 0 \\ \sigma_2 \end{bmatrix}$$

Multiplication

- Visualize: $Ax = U\Sigma V^\top x$

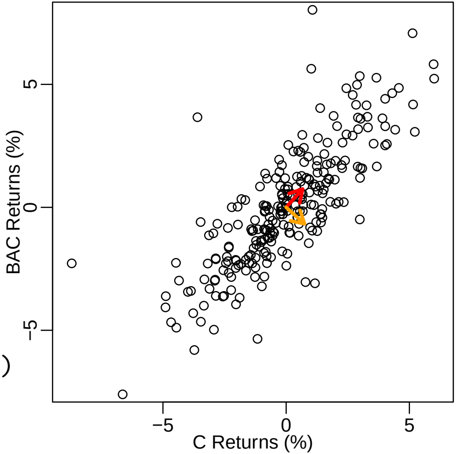
U	rotate right 45°
Σ	stretch x coord by 1.5, stretch y coord by 2
V	rotate right 22.5°



$x := \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$ $\qquad x := V^\top x$ $\qquad x := \Sigma x$ $\qquad x := Ux$

Example

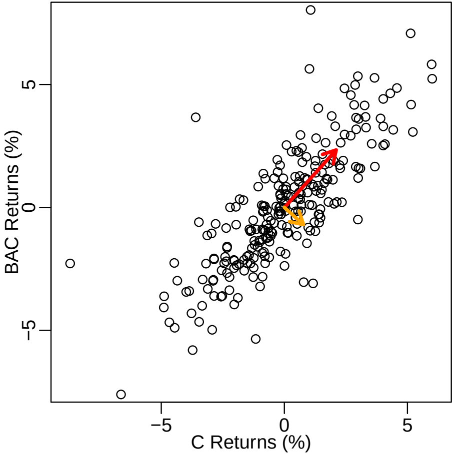
```
> load("R.RData")
> library(MASS)
> eqscplot(R)

> svdR <- svd(R)
> U <- svdR$u
> S <- diag(svdR$d)
> V <- svdR$v

> all.equal(U %*% S %*% t(V), R)
[1] TRUE
```




```
> arrows(0, 0, V[1,1], V[2,1])
> arrows(0, 0, V[1,2], V[2,2])
```

Example (continued)

```
> u <- V[, 1] * S[1, 1]
> u <- u / sqrt(m - 1)
> arrows(0, 0, u[1], u[2])

> w <- V[, 2] * S[2, 2]
> w <- w / sqrt(m - 1)
> arrows(0, 0, w[1], w[2])
```



Outline

Eigenvalues and Eigenvectors

- Let R be an $m \times n$ matrix and let $\tilde{R} = \left(I - \frac{ee^\mathsf{T}}{m}\right) R$
- Let $\tilde{R} = U\Sigma V^\mathsf{T}$ be the singular value factorization of $\tilde{R}$
- Recall that

$$\begin{aligned}
[\mathrm{Cov}(R)] &= \tfrac{1}{m-1}\,\tilde{R}^\mathsf{T}\tilde{R} \\
&= \tfrac{1}{m-1}\,(U\Sigma V^\mathsf{T})^\mathsf{T}(U\Sigma V^\mathsf{T}) \\
&= \tfrac{1}{m-1}\,V\Sigma^\mathsf{T}(U^\mathsf{T}U)\Sigma V^\mathsf{T} \\
&= \tfrac{1}{m-1}\,V\Sigma^\mathsf{T}\Sigma V^\mathsf{T}
\end{aligned}$$

Eigenvalues and Eigenvectors

- Remember that Σ is a diagonal $m \times n$ matrix

$$\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & & \end{bmatrix} \quad \Longrightarrow \quad \Sigma^\mathsf{T}\Sigma = \underbrace{\begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_n & \end{bmatrix}}_{n \times m} \underbrace{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & & \end{bmatrix}}_{m \times n}$$

- $\Sigma^\mathsf{T}\Sigma$ is a diagonal matrix with $\sigma_1^2 \geq \cdots \geq \sigma_n^2$ along the diagonal

- Let

$$\Lambda = \frac{1}{m-1}\Sigma^\mathsf{T}\Sigma = \begin{bmatrix} \lambda_1 = \frac{\sigma_1^2}{m-1} & & \\ & \ddots & \\ & & \lambda_n = \frac{\sigma_n^2}{m-1} \end{bmatrix}$$

Eigenvalues and Eigenvectors

- Substitute Λ into the expression for the covariance matrix of R

$$\big[\mathrm{Cov}(R)\big] = \tfrac{1}{m-1} V \Sigma^{\mathsf{T}} \Sigma V^{\mathsf{T}} = V \Lambda V^{\mathsf{T}}$$

- Let e_j be a unit vector in the j^{th} coordinate direction
- Multiply a right singular vector v_j by $\mathrm{Cov}(R)$

$$\big[\mathrm{Cov}(R)\big] v_j = V \Lambda V^{\mathsf{T}} v_j = V \Lambda \big(V^{\mathsf{T}} v_j\big) = V \Lambda \, e_j$$

- Recall that a matrix times a vector is a linear combination of the columns

$$Av = v_1 \begin{bmatrix} | \\ a_1 \\ | \end{bmatrix} + \cdots + v_1 \begin{bmatrix} | \\ a_1 \\ | \end{bmatrix} \quad \Longrightarrow \quad \Lambda e_j = 1 \begin{bmatrix} \vdots \\ \lambda_j \\ \vdots \end{bmatrix} = \lambda_j e_j$$

Eigenvalues and Eigenvectors

- Substituting $\Lambda e_j = \lambda_j e_j \ldots$

$$[\mathrm{Cov}(R)]\, v_j = V \Lambda V^{\mathsf{T}} v_j = V \Lambda (V^{\mathsf{T}} v_j) = V \Lambda\, e_j = V \lambda_j e_j = \lambda_j V e_j = \lambda_j v_j$$

- In summary
    - v_j is a right singular vector of $\tilde{R}$
    - $[\mathrm{Cov}(R)] = \tilde{R}^{\mathsf{T}} \tilde{R}$
    - $[\mathrm{Cov}(R)]\, v_j = \lambda_j v_j$
    - $[\mathrm{Cov}(R)]\, v_j$ same direction as v_j, length scaled by factor λ_j

- In general: let A be a square matrix and consider the product Ax
    - Certain *special* vectors x are in the same direction as Ax
    - These vectors are called <u>eigenvectors</u>
    - Equation: $Ax = \lambda_x x$; the number λ_x is the <u>eigenvalue</u>

Diagonalizing a Matrix

- Suppose an $n \times n$ matrix A has n linearly independent eigenvectors
- Let S be a matrix whose columns are the n eigenvectors of A

$$S^{-1}AS = \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

- The matrix A is <u>diagonalized</u>
- Useful representations of a diagonalized matrix

$$AS = S\Lambda \qquad S^{-1}AS = \Lambda \qquad A = S\Lambda S^{-1}$$

- Diagonalization requires that A have n eigenvectors
- Side note: invertibility requires nonzero eigenvalues

The Spectral Theorem

- Returning to the motivating example ...
- Let $A = \tilde{R}^\intercal \tilde{R}$ where $\tilde{R}$ is an $m \times n$ matrix
- A is symmetric

- Spectral Theorem Every symmetric matrix $A = A^\intercal$ has the factorization $Q\Lambda Q^\intercal$ with real diagonal Λ and orthogonal matrix Q:

$$A = Q\Lambda Q^{-1} = Q\Lambda Q^\intercal \quad \text{with} \quad Q^{-1} = Q^\intercal$$

- **Caveat** A nonsymmetric matrix can easily produce λ and x that are complex

Positive Definite Matrices

- The symmetric matrix A is positive definite if $x^\mathsf{T} A x > 0$ for every nonzero vector x
- 2×2 case: $x^\mathsf{T} A x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = a x_1^2 + 2b x_1 x_2 + c x_2^2 > 0$
- The scalar value $x^\mathsf{T} A x$ is a quadratic function of x_1 and x_2

$$f(x_1, x_2) = a x_1^2 + 2b\, x_1 x_2 + c x_2^2$$

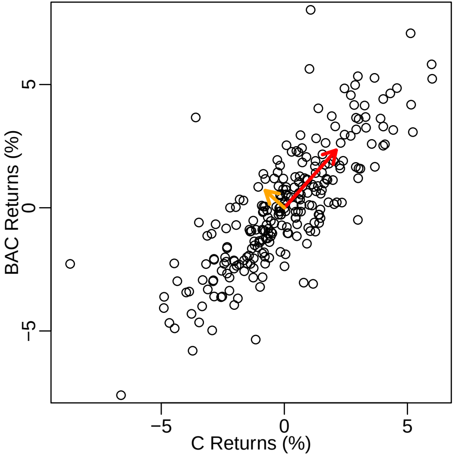
- f has a minimum of 0 at $(0, 0)$ and is positive everywhere else

1×1 a is a positive number

2×2 A is a positive definite matrix

R Example

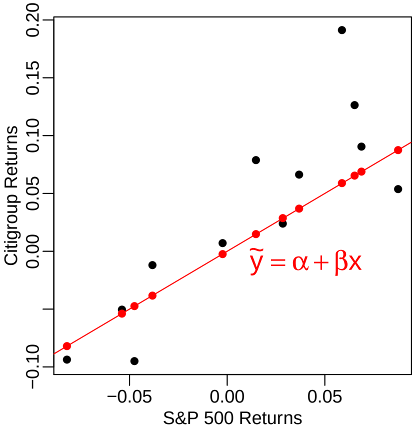
```
> eigR <- eigen(var(R))
> S <- eigR$vectors
> lambda <- eigR$values

> u <- sqrt(lambda[1]) * S[,1]
> arrows(0, 0, u[1], u[2])

> w <- sqrt(lambda[2]) * S[,2]
> arrows(0, 0, w[1], w[2])
```



Outline

Least Squares

Citigroup Returns vs. S&P 500 Returns (Monthly - 2010)



- Set of m points (x_i, y_i)
- Want to find best-fit line

$$\hat{y} = \hat{\alpha} + \hat{\beta}x$$

- Criterion: $\sum_{i=1}^{m} [y_i - \hat{y}_i]^2$ should be minimum
- Choose $\hat{\alpha}$ and $\hat{\beta}$ so that

$$\sum_{i=1}^{m} [y_i - (\alpha + \beta x_i)]^2$$

minimized when

- $\alpha = \hat{\alpha}$
- $\beta = \hat{\beta}$

Least Squares

- What does the column picture look like?
- Let $y = (y_1, y_2, \ldots, y_m)$
- Let $x = (x_1, x_2, \ldots, x_m)$
- Let e be a column vector of m ones
- Can write $\tilde{y}$ as a linear combination

$$\tilde{y} = \begin{bmatrix} \tilde{y}_1 \\ \vdots \\ \tilde{y}_m \end{bmatrix} = \alpha \begin{bmatrix} | \\ e \\ | \end{bmatrix} + \beta \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = X\beta$$

- Want to minimize

$$\sum_{i=1}^{m} \left[y_i - \tilde{y}_i \right]^2 = \|y - \tilde{y}\|^2 = \|y - X\beta\|^2$$

QR Factorization

- Let A be an $m \times n$ matrix with linearly independent columns
- **Full QR Factorization:** A can be written as the product of
  - an $m \times m$ orthogonal matrix Q
  - an $m \times n$ upper triangular matrix R
    (upper triangular means $r_{ij} = 0$ when $i > j$)

$$A = QR$$

- Want to minimize

$$\|y - X\beta\|^2 = \|y - QR\beta\|^2$$

- Recall: orthogonal transformation leaves vector lengths unchanged

$$\|y - X\beta\|^2 = \|y - QR\beta\|^2 = \|Q^\intercal(y - QR\beta)\|^2 = \|Q^\intercal y - R\beta\|^2$$

Least Squares

- Let $u = Q^{\mathsf{T}} y$

$$u - R\beta = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \end{bmatrix} - \begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \\ 0 & 0 \\ \vdots & \vdots \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} u_1 - (r_{11}\alpha + r_{12}\beta) \\ u_2 - r_{22}\beta \\ u_3 \\ \vdots \end{bmatrix}$$

- α and β effect only the first n elements of the vector

- Want to minimize

$$\|u - R\beta\|^2 = \left[u_1 - (r_{11}\alpha + r_{12}\beta)\right]^2 + \left[u_2 - r_{22}\beta\right]^2 + \sum_{i=(n+1)}^{m} u_i^2$$

Least Squares

- Can find $\hat{\alpha}$ and $\hat{\beta}$ by solving the linear system $\tilde{R}\hat{\beta} = \tilde{u}$

$$\begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{bmatrix} \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$\tilde{R}$ first n rows of R, $\tilde{u}$ first n elements of u

- System is already upper triangular, solve using back substitution

R Example

- First, get the data

```
> library(quantmod)
> getSymbols(c("C", "^GSPC"))
> citi <- c(coredata(monthlyReturn(C["2010"])))
> sp500 <- c(coredata(monthlyReturn(GSPC["2010"])))
```

- The x variable is `sp500`, bind a column of ones to get matrix X

```
> X <- cbind(1, sp500)
```

- Compute QR factorization of X and extract the Q and R matrices

```
> qrX <- qr(X)
> Q <- qr.Q(qrX, complete = TRUE)
> R <- qr.R(qrX, complete = TRUE)
```

R Example

- Compute $u = Q^\top y$

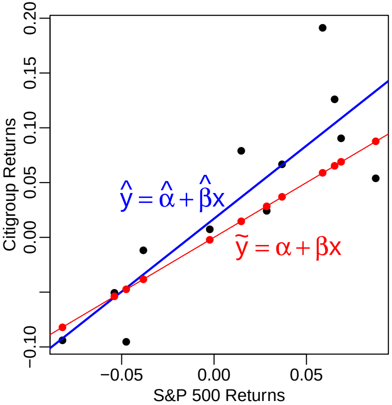
```
> u <- t(Q) %*% citi
```

- Solve for $\hat{\alpha}$ and $\hat{\beta}$

```
> backsolve(R[1:2,1:2], u[1:2])
[1] 0.01708494 1.33208984
```

- Compare with built-in least squares fitting function

```
> coef(lsfit(sp500, citi))
 Intercept          X
0.01708494 1.33208984
```



COMPUTATIONAL FINANCE & RISK MANAGEMENT

UNIVERSITY *of* WASHINGTON

Department of Applied Mathematics

http://computational-finance.uw.edu