Linear Inverse Problems

Monday, March 13, 2017 9:02 AM

$$y = Ax$$

$y \in \mathbb{R}^M$, $A \in \mathbb{R}^{M\times N}$, $x \in \mathbb{R}^N$

observe — given — we want to find x

This is called a linear inverse problem.

We can think of y as containing different indirect observations or measurements

$$y = \begin{bmatrix} \langle x, a_1 \rangle \\ \langle x, a_2 \rangle \\ \vdots \\ \langle x, a_M \rangle \end{bmatrix}$$

where a^T is the m^{th} row of A (a^H if A is complex)

We can have $M > N$, $M = N$, or $M < N$

more observations than unknowns — # of observations = # of unknowns — fewer observations than unknowns.

if $M = N$ and A^{-1} exists then

$$x = A^{-1}y$$

in general, it is never this easy

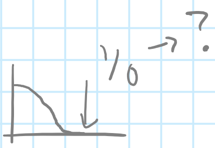
– examples –

Deconvolution – e.g. deblurring, digital communications

$$f(t) \longrightarrow \boxed{\begin{matrix}\text{LTI} \\ h(t)\end{matrix}} \longrightarrow y(t)$$

we know $h(t)$

$$y(t) = \int_{-\infty}^{\infty} h(t-\tau)\, f(\tau)\, d\tau$$

$H(\Omega) =$ 

$Y(\Omega) = H(\Omega)X(\Omega)$

$X(\Omega) = \frac{1}{H(\Omega)} Y(\Omega)$

danger!

↳ more on this later

Linear Inverse applied to Continuous Time

Monday, March 13, 2017 9:22 AM

$$y(t) = \int_{-\infty}^{\infty} h(t-\tau)\, f(\tau)\, d\tau$$

If $f(t)$ is time-limited then we can write

$$f(t) = \sum_n x(n)\, \psi_n(t)$$ where $\{\psi_n(t)\}_n$ forms an orthobasis on a finite interval for $L_2([0,T])$

$$y(t) = \int_{-\infty}^{\infty} h(t-\tau) \sum_n x(n)\, \psi_n(\tau)\, d\tau$$

$$= \sum_n x(n) \underbrace{\int_{-\infty}^{\infty} h(t-\tau)\, \psi_n(\tau)\, d\tau}_{\text{a set of coefficients} \to \text{function of time and of } n.}$$

Generally we observe samples in time of $y(t)$. Suppose we have M observations taken at times

$$t = t_1, t_2, \ldots, t_M$$ these can be arbitrary

$$y[m] \equiv y(t_m) = \sum_n x(n) \left(\int_{-\infty}^{\infty} h(t_m - \tau)\, \psi_n(\tau)\, d\tau \right)$$

$$y[m] = \sum_n A[m,n]\, x[n]$$

where $$A[m,n] = \int_{-\infty}^{\infty} h(t_m - \tau)\, \psi_n(\tau)\, d\tau$$

$$= \langle \underline{h}_m, \underline{\psi}_n \rangle$$

where $$h_m(t) = h(t_m - t)$$

$$\boxed{\underline{y} = A\, \underline{x}}$$

$$\hat{f}(t) = \sum_{n=1}^{N} \hat{x}(n)\, \psi_n(t)$$

Solving Linear Inverse Equations

Monday, March 13, 2017 9:37 AM

We will start with the simplest cases

A is $N \times N$ and symmetric (or Hermitian for complex A)

Definition: If A is real-valued, then we call it **symmetric** if $A^T = A$

($A[m,n] = A[n,m]$ for all $m,n = 1, \ldots, N$)

example $\begin{bmatrix} 1 & 3 & 7 \\ 3 & -5 & -2 \\ 7 & -2 & 6 \end{bmatrix}$

Definition: If A is complex valued, we call it **Hermitian** if $A^H = A$.

$\left(A[m,n] = \overline{A[n,m]} \right)$ example $\begin{bmatrix} 1 & 3+j2 & 1-j3 \\ 3-j3 & -5 & 4 \\ 1+j3 & 4 & -6 \end{bmatrix}$

We will work up to non-symmetric and non-square...

Eigenvalue decompositions of symmetric matrices

Definition: An **eigenvector** of an $N \times N$ matrix is a vector $\mathbf{v}$ such that

$$A\underline{v} = \lambda \underline{v}$$

for some $\lambda \in \mathbb{C}$. The scalar λ is called an **eigenvalue** associated with $\mathbf{v}$

examples: $\underbrace{(A - \lambda I)}\underline{v} = 0$

we can use this to find λ.

$\det(A - \lambda I) = 0$

solve the "characteristic" equation to find λ

matlab

$\downarrow$

eig(A)

Eigenvalue Decompositions

Monday, March 13, 2017 9:51 AM

Suppose that the matrix A has N linearly-independent eigenvectors $\underset{\sim}{v}_1, \ldots, v_N$

$$A v_1 = \lambda_1 v_1$$
$$A v_2 = \lambda_2 v_2$$
$$\vdots \qquad \vdots$$
$$A v_N = \lambda_N v_N$$

$$A \underbrace{\begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_N \\ | & | & & | \end{bmatrix}}_{V} = \underbrace{\begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_N \\ | & | & & | \end{bmatrix}}_{V} \underbrace{\begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{bmatrix}}_{\Lambda}$$

$$AV = V\Lambda$$ if the $\underset{\sim}{v}_i$'s are linearly independent

$$\boxed{A = V \Lambda V^{-1}} \iff \boxed{\Lambda = V^{-1} A V}$$

This does not work for all matrices! But, it does work for all symmetric matrices.

↳ comes from the Schur Triangularity Lemma – any $N \times N$ matrix is uniterily similar to an upper-triangular matrix. That is, for a given $B \in \mathbb{C}^{N \times N}$ there is an orthormal matrix V such that

$$B = V \Delta_u V^H$$

where

$$\Delta = \begin{bmatrix} \Delta[1,1] & \Delta[1,2] & \cdots & \Delta[1,N] \\ & \Delta[2,2] & & \vdots \\ & 0 & \ddots & \\ & & & \Delta[N,N] \end{bmatrix}$$

Diagonalization

Monday, March 13, 2017 9:51 AM

If B is Hermitian the $B = B^H$ implies

$$V \Delta V^H = (V \Delta V^H)^H = V \Delta^H V^H$$

so $\Delta = \Delta^H$ -> since it is upper triangular, it must be diagonal and real.

$$\Delta = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots & \\ & & & \lambda_N \end{bmatrix}, \quad \lambda_n \in \mathbb{R}$$

if $B \in \mathbb{R}^{N \times N}$ and symmetric, it is also Hermitian so any real, symmetric or complex Hermitian matrix is diagonalizable.

Note that $V^H = V^{-1}$ — The eigenvector matrices contain ortho. columns.

Recall $A \underline{v}_n = \lambda_n \underline{v}_n$ => let $\underline{u}_n = \alpha \underline{v}_n$

then $A u_n = \alpha A v_n = \alpha \lambda_n v_n = \lambda_n u_n$

thus, we can normalize our eigenvectors

Other properties

An $N \times N$ symmetric/Hermitian matrix A has:

- Real eigenvalues (even if A is complex) $\lambda_1, \ldots, \lambda_N$
- N orthogonal eigenvectors. $\underline{v}_1, \ldots, \underline{v}_N$
- If $A \in \mathbb{R}^{N \times N}$, Then $\underline{v}_n$ can be chosen to be real-valued

$$A = V \Lambda V^H = \sum_{n=1}^{N} \lambda_n \underline{v}_n \underline{v}_n^H$$

Matrix Spectral Decomposition

Monday, March 13, 2017 10:19 AM

$$A = \underbrace{\begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_N \\ | & | & & | \end{bmatrix}}_{V} \underbrace{\begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} — & v_1^H & — \\ — & v_2^H & — \\ & \vdots & \\ — & v_N^H & — \end{bmatrix}}_{V^H}$$

A is a sum of outer products $(v_n v_n^H)$'s weighted by λ_n's

$$A = \sum_{n=1}^{N} \lambda_n v_n v_n^H = \sum_{n=1}^{N} \lambda_n v_n v_n^T$$

if A real

This is often called the spectral decomposition of A

Schur decomposition notes

Thursday, March 16, 2017 9:40 AM

Technical details: Schur decomposition

In this section we prove one of the fundamental results in linear algebra: that any $N \times N$ matrix is *unitarily similar* to an upper-triangular matrix. That is, given an $N \times N$ matrix $\boldsymbol{A}$, there is an orthonormal matrix $\boldsymbol{V}$ (meaning $\boldsymbol{V}^{\mathrm{H}}\boldsymbol{V} = \mathbf{I}$) such that

$$\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Delta}\boldsymbol{V}^{\mathrm{H}},$$

where

$$\boldsymbol{\Delta} = \begin{bmatrix} \Delta[1,1] & \Delta[1,2] & \cdots & \Delta[1,N] \\ 0 & \Delta[2,2] & \cdots & \Delta[2,N] \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \Delta[N,N] \end{bmatrix}.$$

This is known as the **Schur Decomposition** or the **Schur Triangulation**. It is also possible to choose $\boldsymbol{V}$ so that $\boldsymbol{\Delta}$ is lower-triangular.

The proof works by induction. First, we use the fact that every matrix has at least one eigenvector. Let $\boldsymbol{v}_1$ be an eigenvector of $\boldsymbol{A}$; we may assume that $\boldsymbol{v}_1$ is normalized, since all scalar multiples of eigenvectors are also eigenvectors. Then we take $\boldsymbol{V}_1$ to be any orthogonal matrix with $\boldsymbol{v}_1$ as one of its columns:

$$\boldsymbol{V}_1 = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{U}_1 \end{bmatrix}, \quad \boldsymbol{U}_1 \in \mathbb{R}^{N \times N-1}, \quad \boldsymbol{U}_1^{\mathrm{H}}\boldsymbol{U}_1 = \mathbf{I}, \quad \boldsymbol{U}_1^{\mathrm{H}}\boldsymbol{v}_1 = \mathbf{0}.$$

This is equivalent to finding an orthobasis for $\mathbb{R}^N$ where $\boldsymbol{v}_1$ is one of the basis vectors and the $N-1$ columns of $\boldsymbol{U}_1$ are the others. There are many such choices for $\boldsymbol{U}_1$; one can be found using the Gram-Schmidt algorithm.

Since $\boldsymbol{v}_1$ is an eigenvector of $\boldsymbol{A}$ (call the corresponding eigenvalue λ_1),

$$\boldsymbol{A}\boldsymbol{V}_1 = \begin{bmatrix} \lambda_1\boldsymbol{v}_1 & \boldsymbol{A}\boldsymbol{U}_1 \end{bmatrix},$$

and

$$\boldsymbol{V}_1^{\mathrm{H}}\boldsymbol{A}\boldsymbol{V}_1 = \begin{bmatrix} \lambda_1 & \\ 0 & \\ \vdots & \boldsymbol{V}_1^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}_1 \\ \vdots & \\ 0 & \end{bmatrix}.$$

Now suppose we have an $N \times N$ matrix of the form

$$\boldsymbol{A}_p = \begin{bmatrix} \boldsymbol{\Delta}_p & \boldsymbol{W}_p \\ \mathbf{0} & \boldsymbol{M}_p \end{bmatrix}, \tag{1}$$

where $\boldsymbol{\Delta}_p$ is a $p \times p$ upper-triangular matrix, $\boldsymbol{W}_p$ is an arbitrary $p \times (N-p)$ matrix, and $\boldsymbol{M}_p$ is an arbitrary $(N-p)\times(N-p)$ square matrix. Now let $\boldsymbol{v}_{p+1}$ be an eigenvector of $\boldsymbol{M}_p$ with corresponding eigenvalue λ_{p+1}, and let $\boldsymbol{U}_{p+1}$ be an $(N-p) \times (N-p-1)$ matrix such that

$$\boldsymbol{Z}_{p+1} = \begin{bmatrix} \boldsymbol{v}_{p+1} & \boldsymbol{U}_{p+1} \end{bmatrix}$$

is a $(N-p) \times (N-p)$ orthonormal matrix. Set

$$\boldsymbol{V}_{p+1} = \begin{bmatrix} \mathbf{I}_p & \mathbf{0} \\ \mathbf{0} & \boldsymbol{Z}_{p+1} \end{bmatrix},$$

where $\mathbf{I}_p$ is the $p \times p$ identity matrix. It should be clear that $\boldsymbol{V}_{p+1}$ is an orthonormal matrix. Applying $\boldsymbol{V}_{p+1}$ to the right of $\boldsymbol{A}_p$ yields

$$\boldsymbol{A}_p\boldsymbol{V}_{p+1} = \begin{bmatrix} \boldsymbol{\Delta}_p & \boldsymbol{W}_p\boldsymbol{Z}_{p+1} \\ \mathbf{0} & [\lambda_{p+1}\boldsymbol{v}_{p+1} \;\; \boldsymbol{M}_p\boldsymbol{U}_{p+1}] \end{bmatrix},$$

and so

$$\boldsymbol{V}_{p+1}^{\mathrm{H}}\boldsymbol{A}_p\boldsymbol{V}_{p+1} = \begin{bmatrix} \boldsymbol{\Delta}_p & \boldsymbol{W}_p\boldsymbol{Z}_{p+1} \\ \mathbf{0} & \begin{bmatrix} \lambda_{p+1} & \\ 0 & \\ \vdots & \boldsymbol{Z}_{p+1}^{\mathrm{H}}\boldsymbol{M}_p\boldsymbol{U}_{p+1} \\ \vdots & \\ 0 & \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Delta}_{p+1} & \boldsymbol{W}_{p+1} \\ \mathbf{0} & \boldsymbol{M}_{p+1} \end{bmatrix},$$

Schur decomposition notes

Thursday, March 16, 2017 9:40 AM

where $\mathbf{\Delta}_{p+1}$ is a $(p+1) \times (p+1)$ upper-triangular matrix, and $\boldsymbol{W}_{p+1}$ and $\boldsymbol{M}_{p+1}$ are arbitrary $(p+1) \times (N-p-1)$ and $(N-p-1) \times (N-p-1)$ matrices, respectively.

Given an arbitrary $\boldsymbol{A}$,

$$\boldsymbol{A}_p = \boldsymbol{V}_{p-1}^{\mathrm{H}} \cdots \boldsymbol{V}_2^{\mathrm{H}} \boldsymbol{V}_1^{\mathrm{H}} \boldsymbol{A} \boldsymbol{V}_1 \boldsymbol{V}_2 \cdots \boldsymbol{V}_{p-1}$$

will have the form (1). Applying the construction over N iterations gives

$$\mathbf{\Delta} = \boldsymbol{V}_N^{\mathrm{H}} \cdots \boldsymbol{V}_2^{\mathrm{H}} \boldsymbol{V}_1^{\mathrm{H}} \boldsymbol{A} \boldsymbol{V}_1 \boldsymbol{V}_2 \cdots \boldsymbol{V}_N,$$

which will be upper-triangular. Since each of the $\boldsymbol{V}_p$ are orthonormal, $\boldsymbol{V} := \boldsymbol{V}_1 \boldsymbol{V}_2 \cdots \boldsymbol{V}_N$ will also be orthonormal. Thus

$$\mathbf{\Delta} = \boldsymbol{V}^{\mathrm{H}} \boldsymbol{A} \boldsymbol{V} \quad \Leftrightarrow \quad \boldsymbol{A} = \boldsymbol{V} \mathbf{\Delta} \boldsymbol{V}^{\mathrm{H}},$$

where $\mathbf{\Delta}$ is upper-triangular and $\boldsymbol{V}^{\mathrm{H}} \boldsymbol{V} = \mathbf{I}$.

Eigenvalues of $\boldsymbol{A}$

The diagonal entries of the matrix $\mathbf{\Delta}$ will contain the λ_p used in the construction above (which we might recall are the eigenvalues of the submatrices $\boldsymbol{M}_p$):

$$\Delta[p,p] = \lambda_p.$$

We can see now that the λ_p are also eigenvalues of $\boldsymbol{A}$. Since $\mathbf{\Delta}$ is triangular, its diagonal entries $\lambda_1, \ldots, \lambda_N$ are its eigenvalues. If $\boldsymbol{x}_p$ is the eigenvector of $\mathbf{\Delta}$ corresponding to λ_p, then taking $\boldsymbol{y}_p = \boldsymbol{V}\boldsymbol{x}_p$ we have

$$\boldsymbol{A}\boldsymbol{y}_p = \boldsymbol{V}\mathbf{\Delta}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{V}\boldsymbol{x}_p = \boldsymbol{V}\mathbf{\Delta}\boldsymbol{x}_p = \lambda_p \boldsymbol{V}\boldsymbol{x}_p = \lambda_p \boldsymbol{y}_p,$$

and so the $\lambda_1, \ldots, \lambda_N$ are eigenvalues of $\boldsymbol{A}$ as well.

Schur decomposition notes

Thursday, March 16, 2017 9:40 AM

Real-valued decompositions

If $\boldsymbol{A}$ is real-valued but non-symmetric, then both $\boldsymbol{V}$ and $\boldsymbol{\Delta}$ can be complex-valued. However, there does real-valued $\boldsymbol{U}$ and $\boldsymbol{\Upsilon}$ such that

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Upsilon}\boldsymbol{U}^{\mathrm{T}},$$

where $\boldsymbol{U}$ is orthonormal, $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U} = \mathbf{I}$, and Υ is almost upper-triangular:

$$\boldsymbol{\Upsilon} = \begin{bmatrix} \boldsymbol{\Lambda}_1 & * & \cdots & * \\ 0 & \boldsymbol{\Lambda}_2 & \cdots & * \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \boldsymbol{\Lambda}_K \end{bmatrix}.$$

The $\boldsymbol{\Lambda}_p$ above are either 2×2 matrices or scalars; there is a 2×2 block for every pair of complex-conjugate eigenvalues of $\boldsymbol{A}$, and a scalar for every real eigenvalue. Although this decomposition is not strictly upper-triangular, it carries many of the same advantages. For example, with $\boldsymbol{U}$ pre-computed and given a $\boldsymbol{b} \in \mathbb{R}^N$, we can still compute the solution to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ with $O(N^2)$ operations.

Symmetric PD Matrices cont.

Monday, March 20, 2017 9:02 AM

In-class attendance quiz.

Find the eigenvalues of $\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$

Hint: $(A - I\lambda)v = 0$

← so (this) is singular

$$\det(A-\lambda I) = (3-\lambda)(2-\lambda) - 1 = 0$$

$$6 - 5\lambda + \lambda^2 - 1 = 0$$

$$\lambda^2 - 5\lambda + 5 = 0$$

$$\lambda = \frac{5 \pm \sqrt{25-20}}{2}$$

$$\lambda = \frac{5}{2} \pm \frac{\sqrt{5}}{2}$$

How to find eigenvectors?

$$Av_1 = \lambda_1 v_1$$

$$\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix} = \lambda_1 \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix}$$ ← solve 2 eq., 2 unknowns

or use eig(A) in matlab

An $N \times N$ symmetric/Hermitian matrix A has:

- Real eigenvalues, $\lambda_1, \dots, \lambda_N$ (even for complex A)
- N orthogonal eigenvectors, $v_1, \dots, v_N$
- If A is real-valued, then v_n can be chosen to be real-valued

We can decompose real-valued A as

$$A = V \Lambda V^T = \sum_{n=1}^{N} \lambda_n v_n v_n^T$$

eigenvectors as columns

$Ax = \text{(inverse } V \text{ transform)}$ ③ $\text{(pointwise multiply)}$ ② $\cdot \underline{\underline{(V \text{ transform})\, x}}$ ①

Symmetric PD matrices cont.

Monday, March 20, 2017 9:28 AM

Definition: a symmetric matrix A is called positive definite if it has positive eigenvalues

PD

$$\lambda_n > 0 \quad \text{for } n = 1, \dots, N$$

we call it positive semi-definite if $\lambda_n \geq 0$ for $n = 1, \dots, N$

We typically assume that

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_N$$

Example

$$A = \frac{1}{2}\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \Rightarrow v_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

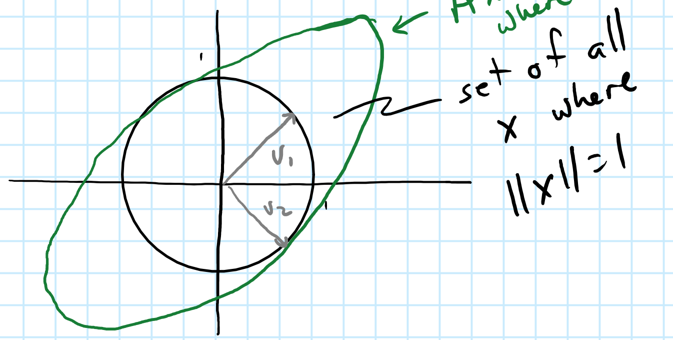
$$\lambda_1 = 2 \qquad \lambda_2 = 1$$

$$Ax = V\Lambda V^T x \qquad \text{suppose } x = \alpha_1 v_1 + \alpha_2 v_2$$

$$= \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix} (\alpha_1 v_1 + \alpha_2 v_2)$$

$$= \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} 2\alpha_1 \\ \alpha_2 \end{bmatrix}$$

$$= 2\alpha_1 v_1 + \alpha_2 v_2$$



Symmetric Positive Definite Matrices

Monday, March 20, 2017 9:42 AM

for sym+def (symmetric positive definite) A,

$$\max_{x \in \mathbb{R}^N} \frac{x^T A x}{\|x\|_2^2} = \max_{\substack{x \in \mathbb{R}^N \\ \|x\|_2 = 1}} x^T A x = \lambda_1 \quad \text{largest eigenvalue}$$

$$\min_{x \in \mathbb{R}^N} \frac{x^T A x}{\|x\|_2^2} = \min_{\substack{x \in \mathbb{R}^N \\ \|x\|_2 = 1}} x^T A x = \lambda_N \quad \text{smallest eigenvalue}$$

$$\max_{\substack{x \in \mathbb{R}^N \\ \|x\|_2 = 1}} x^T A x = \max_{\cdots} x^T V \Lambda V^T x = \max_{\substack{y \in \mathbb{R}^N \\ \|y\|=1}} y^T \Lambda y = \max_{\|y\|_2 = 1} \sum (y[n])^2 \lambda_n$$

where $y = V^T x$

maxi'zed when

$$\begin{cases} y[1] = 1, \\ y[k] = 0 \text{ for } k \neq 1 \end{cases}$$

solving equations

If $y = Ax$ and A is sym+def., then

$$x = A^{-1} y$$

$$x = V \Lambda^{-1} V^T y$$

$$x = \sum_{n=1}^{N} \frac{1}{\lambda_n} \langle y, v_n \rangle v_n$$

once we have the λ_n's and v_n's,

The solution becomes a matrix multiply and sum of vectors.

Eigenvalues/eigenvectors of A^{-1}?

if A is sym+def then

$$A v_n = \lambda_n v_n = v_n = \lambda_n A^{-1} v_n$$

$$\boxed{A^{-1} v_n = \frac{1}{\lambda_n} v_n}$$

The eigenvalues are $\left\{ \frac{1}{\lambda_N}, \frac{1}{\lambda_{N-1}}, \ldots, \frac{1}{\lambda_1} \right\}$

eigenvectors are $v_N, \ldots, v_1$

Observation Error

Monday, March 20, 2017 9:57 AM

Now suppose we have some observation error

$$\underline{y} = A\underline{x} + \underline{e}$$

unknown error vector in $\mathbb{R}^N$

$$\tilde{x} = A^{-1}y = A^{-1}(Ax + e) = x + A^{-1}e$$

the error in our solution is

$$\boxed{\underline{\tilde{x}} - \underline{x} = A^{-1}\underline{e}}$$

suppose $\|e\|_2 = 0.001$, what does $A^{-1}\underline{e}$ look like?

what is the largest eigenvalue of A^{-1}? $1/\lambda_N$

The error could be amplified by as much as $1/\lambda_N$

If $\lambda_N = 10^{-5}$, the error in $\tilde{x}$ is (for our example)

$\boxed{100\,\underline{e}}$

example, A blurs rows of an image. Then there is some noise added from the sensor

blurring operation  ← picture with shaky camera

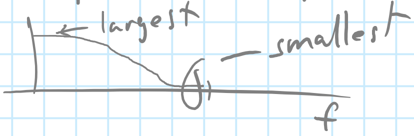
$y = Ax + e$

image pixel row

pixel error from CMOS sensor

a blurring has a lowpass effect

so A^{-1} will amplify the noise...



note, as the dimensionality of A increases, the smallest and largest eigenvalues tend toward the spectral min & max respectively. (more details later)

Reconstruction Error

Monday, March 20, 2017 10:12 AM

$$\frac{1}{\lambda_1^2}\|e\|_2^2 \le \|\tilde{x}-x\|_2^2 \le \frac{1}{\lambda_N^2}\|e\|_2^2$$

best case

worst case

average reconstruction error?

we will assume that the entries e are random, iid Gaussian

$$e[n] \sim \text{Normal}(0,\sigma^2) \qquad n=1,\dots,N$$

$$E\{e[n]e[\ell]\} = \begin{cases}\sigma^2 & n=\ell\\ 0 & n\neq\ell\end{cases}$$

$$E\{\|e\|_2^2\} = E\left\{\sum_{n=1}^N |e[n]|^2\right\}$$

$$= \sum_{n=1}^N E\{e^2[n]\} = \sum_{n=1}^N \sigma^2$$

$$E\{\|e\|_2^2\} = N\sigma^2$$

$$E\{\|A^{-1}e\|_2^2\} = E\{\langle V\Lambda^{-1}V^Te, V\Lambda^{-1}V^Te\rangle\}$$

$$(V\Lambda^{-1}V^Te)^T(V\Lambda^{-1}V^Te)$$

$$(e^TV\Lambda^{-1})\cancel{V^T}\cancel{V}(\Lambda^{-1}V^Te)$$

$$= E\{\langle \Lambda^{-1}V^Te, \underbrace{V^TV}_{I}\Lambda^{-1}V^Te\rangle\}$$

$$= E\left\{\sum_{n=1}^N \left|\frac{1}{\lambda_N}\langle e,v_n\rangle\right|^2\right\}$$

$$= \sum_{n=1}^N \frac{1}{\lambda_N^2} E\{|\langle e,v_n\rangle|\} \quad -\sigma^2 \quad = N\sigma^2\left(\underbrace{\frac{1}{N}\sum_n \frac{1}{\lambda_n^2}}\right)$$

average eigenvalue squared of A^{-1}

$$\left(\sum_{m=1}^N v_n[m]e[m]\right)^2$$

$$= \sum_{m=1}^N\sum_{\ell=1}^N v_n[m]e[m]v_n[\ell]e[\ell]$$

0 except m=l

$$E\{(\;\cdot\;)\} = \sum_m\sum_\ell v_n[m]v_n[\ell]E\{e[m]e[\ell]\}$$

$$= \sum_m |v_n[m]|^2\sigma^2 = \sigma^2$$