SINGULAR VALUE DECOMPOSITION, EOF'S AND INVERSES

Decomposition of data into Fourier components is particularly useful for stationary processes (using the terminology of time series) because (a) the Fourier functions are orthogonal over the length of the time series and (b) the amplitudes of different modes (frequencies) are uncorrelated when averaged over many series. It is perfectly straightforward to find analogous orthogonal and uncorrelated functions to describe any process regardless of the degree of nonstationarity. These functions go by various names in different fields; in oceanography and meteorology they are called Empirical Orthogonal Functions. The two basic reasons why EOFs are useful are (a) EOFs provide the most accurate possible representation of the field in terms of a limited number of functions and (b), because their amplitudes are uncorrelated, EOFs are logically distinct modes of variability.

Finding EOFs can be regarded as selecting, for a particular data set, a more appropriate set of functions than the sines and cosines of Fourier analysis. In this context we ask for functions that efficiently describe energy averaged over an ensemble of realizations. EOFs are closely related to the **singular value decomposition (SVD)** of a matrix that arises in many algebraic applications: linear estimation, solution of differential equations, inverse modeling, etc. It is generally useful in analyzing and understanding linear transformations. The many applications of the SVD provide motivation for understanding it. One of the clearest explanations is provided by Lanczos (*Linear Differential Operators*, chapter 3), where the SVD is derived as an extension of eigenvalue analysis.

To take a concrete example, suppose we have a time series of maps $A(x,t)$ which cover N_T time values t and N_X positions x. The observations then form the matrix $\mathbf{A}$. Following the rules of linear algebra, in a linear vector space it is always possible to describe each of the maps using a set of N_X functions of x if these functions form a complete basis for the x-space.

$$A(x,t) = \sum_{k}^{N} U(x,k)V^*(t,k) \qquad \text{or} \qquad \mathbf{A} = \mathbf{U}\cdot\mathbf{V}^T \tag{1}$$

The superscript T denotes the conjugate transpose (also called the Hermitian). In Fourier analysis $U(x,k)$ would be the complex exponentials with different wavenumbers indicated by the index k. $V(t,k)$ would be the Fourier amplitude for that wavenumber for series t.

Following the earlier discussion of linear models, it is easiest to insure that $\mathbf{U}$ is a complete basis is to make the different functions (columns of $\mathbf{U}$) orthogonal and to insist there be N_X of them.

$$\sum_{t}^{N_X} U(x,n)U^*(x,m) = \delta_{nm} \qquad \text{or} \qquad \mathbf{U}^T\cdot\mathbf{U} = \mathbf{I} \tag{2}$$

Since any complete basis can be rotated to make it orthonormal, this imposes no restriction. Orthonormality make $\mathbf{U}^T$ the inverse of $\mathbf{U}$ so that we can find the amplitudes easily:

$$V(t,k) = \sum_{x}^{N} A^*(x,t)U(x,k) \qquad \text{or} \qquad \mathbf{A}^T\cdot\mathbf{U} = (\mathbf{U}\cdot\mathbf{V}^T)^T\cdot\mathbf{U} = \mathbf{V}\cdot(\mathbf{U}\cdot\mathbf{U}^T) = \mathbf{V} \tag{3}$$

Recall we have earlier shown that if columns of $\mathbf{U}$ are orthogonal, so are the rows,

$$\mathbf{U}^T\cdot\mathbf{U} = \mathbf{I} \qquad \text{and} \qquad \mathbf{U}\cdot\mathbf{U}^T = \mathbf{I}\,, \tag{4}$$

so that $\mathbf{U}$ and $\mathbf{U}^T$ are inverse pairs taken either way. In the Fourier case $\mathbf{U}^T$ is the inverse transform which involves the conjugates of the same complex exponentials as appear in the transform.

1. Empirical Orthogonal Functions (EOFs)

EOFs are useful in describing data because they are the most efficient way to describe the N_X-point maps using less than N_X modes. (Any set of N_X linearly independent functions would work perfectly). More precisely, when $L \;\; N_X$ and the $U(x,k)$ are EOFs, the approximation

$$\hat{A}(x,t) = \sum_{k=1}^{L} U(x,k)\hat{V}^*(t,k) \qquad \text{or} \qquad \hat{\mathbf{A}} = \mathbf{U}\cdot\hat{\mathbf{V}}^T, \tag{5}$$

has the smallest mean square error

$$\langle E\rangle = \frac{1}{N_T}\sum_{t}^{N_T} E(t) \qquad \text{where} \qquad E(t) \equiv \sum_{x}^{N_X} |A(x,t) - \hat{A}(x,t)|^2 \tag{6}$$

possible for an L-function approximation.

To show the efficiency of EOFs we must consider all families of L functions. These functions will not all be orthogonal. But any set of L functions can be rotated (an infinite number of different ways) to make a new set of orthonormal functions without changing their ability to describe any function. The important thing is, then, not what the functions are (which can be changed by rotations and scaling) but what subspace they span. Therefore we can, without loss of generality, consider only functions $U(x,t)$ which for $k \le L$ are orthogonal to each other.

Having selected a set of functions $U(x,k)$ for $k \le L$, or more exactly the subspace that they span, we have also specified the subspace of the original N_X-dimension x-space that these functions do not span. This is the space (of dimension $N_X - L$) of functions that cannot be generated by the selected L functions because they are orthogonal to the selected function. While we do not intend to use them, we note that this space is spanned by $N_X - L$ functions that are orthogonal to the first L functions and, as above, we can take as mutually orthogonal. We will denote these functions as $U(x,k)$ for $k > L$.

Thus any selection of L functions is equivalent to selecting the first L functions $U(x,k)$ of a unitary, N_X by N_X basis $\mathbf{U}$ that obeys (2) and (4). Because the columns of $\mathbf{U}$ are orthogonal, the error of the L-term approximation, defined in (6), is

$$E(t) = \sum_{k=1}^{L} |\hat{V}(t,k) - V(t,k)|^2 + \sum_{k>L}^{N} |V(t,k)|^2 \tag{7}$$

where the $V(t,k)$ is the amplitude from (3) that goes with function $U(x,k)$ when all N_X functions are used. Clearly the best choice is $\hat{V}(t,k) = V(t,k)$. Thus $E(t)$ is simply the sum of the squared amplitudes of those functions not available among the L chosen functions.

Using (3) to find the amplitudes for $k > L$ gives

$$\langle E\rangle = \sum_{k>L}\frac{1}{N_T}\sum_{t}\left[\sum_{x}\sum_{y} U(x,k)A(x,t)A^*(y,t)U^*(y,k)\right] \tag{8}$$

This shows how the data $A(x,t)$ enters into the problem through

$$C(x,y) \equiv \frac{1}{N_T}\sum_{t} A(x,t)A^*(y,t) \qquad \text{or} \qquad \mathbf{C} \equiv \frac{1}{N_T}\mathbf{A}\cdot\mathbf{A}^T \tag{9}$$

which will be recognized as the two-position mean product of A; if we had removed the mean A it would be the two-position covariance of $\langle A\rangle$. In analysis of stationary time series this is would be to taken to depend only on $x - y$ but here we keep the general case.

We are now interested in finding the special functions $U(x,k)$ that make the L-term approximation (5) have the smallest mean error as given in (8). To do this we want to minimize (8) while maintaining the size of each $U(x,k)$. In terms of $\mathbf{C}$, we want to minimize

$$\langle E\rangle = \sum_{k>L}\left[\sum_x\sum_y U(x,k)C(x,y)U^*(y,k) \;-\; \lambda_k^2\sum_x U^2(x,k) - 1\right] \tag{10}$$

or

$$\langle E\rangle = \sum_{k>L}[\mathbf{u}_k^T\cdot\mathbf{C}\cdot\mathbf{u}_k - \lambda_k^2\mathbf{u}_k^T\cdot\mathbf{u}_k] \tag{11}$$

where $\mathbf{u}_k$ is the $k^{\rm th}$ column of $U(x,k)$. The first term is the error and the second term, involving the Lagrange multiplier λ_k^2 is zero when we require $\mathbf{U}$ to obey (2). As the $U(x,k)$ are adjusted to minimize $\langle E\rangle$, these extra terms are used to simplify enforcing the constraint (2) as is common in the calculus of variation. If this constraint were not satisfied then (7), and hence (8) and (10), would not correctly give the error.

To choose the $U(x,k)$ that minimize $\langle E\rangle$ we extremize (10) with respect to variations of each element of $U(x,k)$ for $k>L$ and then adjust the Lagrange multiplier to satisfy the orthonormality conditions (2). Since $\langle E\rangle$ is quadratic in these elements and positive semi-definite, any extremum is a minimum. For each k, differentiating with respect to each element of $U(x,k)$ gives $2\cdot N_X$ equations for the $2\cdot N_X$ (real and imaginary parts) unknowns. A little algebra shows these equations to have the form

$$\sum_y^{N_X} C(x,y)U(y,k) - \lambda_k^2 U(x,k) = 0 \quad \text{or} \quad \mathbf{C}\cdot\mathbf{u}_k - \lambda_k^2\mathbf{u}_k = 0 \tag{12}$$

Thus the most efficient functions are the eigenvectors of the covariance-like matrix $\mathbf{C}=\mathbf{A}\cdot\mathbf{A}^T$. When $\mathbf{C}$ is a plausible approximation of the two-time covariance, the eigenvectors, $\mathbf{u}_k$, are the Empirical Orthogonal Functions and the representation describes variance in the most efficient possible way.

The "amplitudes" $V(t,k)$ of the modes $\mathbf{u}_k$ are automatically orthogonal when the modes are the eigenvectors of (11). To see this, use (3) to find the amplitudes of two different modes and correlate them

$$\frac{1}{N_T}\sum_t V(m,j)V^*(m,k) = \sum_x\sum_y U(x,j)C(x,y)U^*(y,k) = \lambda_k^2\sum_x U(x,j)U^*(x,k) = \lambda_k^2\delta_{t\tau} \tag{13}$$

or

$$\frac{1}{N_T}\mathbf{v}_k^T\cdot\mathbf{v}_j \;=\; \frac{1}{N_T}\mathbf{u}_k^T\cdot\mathbf{A}\cdot\mathbf{A}^T\cdot\mathbf{u}_k \;=\; \mathbf{u}_k^T\mathbf{C}\cdot\mathbf{u}_j \;=\; \lambda_j^2\mathbf{u}_k\cdot\mathbf{u}_j \;=\; \lambda_j^2\delta_{jk} \tag{14}$$

where $\mathbf{v}_k$ is the k^{th} column of $V(m,k)$. This shows that the $\mathbf{v}_k$ are orthogonal and that the eigenvector λ_k^2 is the "mean square" of amplitude $\mathbf{v}_k$, that is $(\mathbf{v}_k^T\cdot\mathbf{v}_k)/N_T \;=\; \lambda_k^2$.

2. SVD, EOFs and Fourier spectra

We have shown that any matrix can be represented as $\mathbf{A}=\mathbf{U}\cdot\mathbf{V}^T$ where the functions $\mathbf{u}_k$ and amplitudes $\mathbf{v}_k$ are orthogonal. We have also shown that the eigenvectors $\mathbf{u}_k$ of (11) are the most efficient orthogonal modes to describe the matrix $\mathbf{A}$ using fewer functions than the rank of A. The Singular Value Decomposition of $\mathbf{A}$ is written in a slightly different way in terms of the

diagonal matrix Λ with $\Lambda_{kk} = \lambda_k$. For a diagonal matrix the inverse Λ^{-1} is a diagonal matrix with diagonal elements $1/\lambda_k$ and $\Lambda^2 = \Lambda \cdot \Lambda$ is diagonal with elements λ_k^2. For the SVD of $\mathbf{A}$ define $\tilde{\mathbf{V}} \equiv \Lambda^{-1} \cdot \mathbf{V}$ so that

$$\mathbf{A} = \mathbf{U} \cdot \Lambda \cdot \tilde{\mathbf{V}}^T \qquad \text{where} \qquad \mathbf{U} \cdot \mathbf{U}^T = \mathbf{U}^T \cdot \mathbf{U} \qquad \text{and} \qquad \tilde{\mathbf{V}} \cdot \tilde{\mathbf{V}}^T = \tilde{\mathbf{V}}^T \cdot \tilde{\mathbf{V}} \textbf{??} \tag{15}$$

The analog of the defining relation (11) is

$$\tilde{C} \equiv \mathbf{A}^T \cdot \mathbf{A} \qquad \tilde{\mathbf{C}} \cdot \mathbf{U} - \mathbf{U} \cdot \Lambda^2 = 0 \qquad \text{or} \qquad (\mathbf{A} \cdot \mathbf{A}^T) \cdot \mathbf{U} - \mathbf{U} \cdot \Lambda^2 = 0 \tag{16}$$

There is nothing special about either variable t or x. This is shown by pre-multiplying (15) by $\mathbf{A}^T$ to give

$$\mathbf{A}^T \cdot \mathbf{A} \cdot (\mathbf{U} \cdot \mathbf{\Lambda} \cdot \mathbf{V}^T)^T \cdot \mathbf{U} - (\mathbf{U} \cdot \Lambda \cdot \mathbf{V}^T)^T \cdot \mathbf{U} \cdot \Lambda^2 = (\mathbf{A}^T \cdot \mathbf{A}) \cdot \mathbf{V} \cdot \Lambda - \mathbf{V} \cdot \Lambda \cdot \Lambda^2 \tag{17}$$

Factoring out Λ yields an eigen-problem which is complementary to (15):

$$(\mathbf{A}^T \cdot \mathbf{A}) \cdot \mathbf{V} \; - \; \mathbf{V} \cdot \Lambda^2 \qquad \text{or} \qquad (\mathbf{A}^T \cdot \mathbf{A}) \cdot \mathbf{v}_k \; - \; \lambda_k^2 \mathbf{v}_k \; = \; 0 \tag{18}$$

Regardless of which dimension you think of being described by EOFs, the appropriate functions come from the same problem. In fact, you can solve either eigenproblem (15) or (16). The smart choice is obvious. Since solving the eigenvalue problem is computationally much more difficult than finding the second set of functions using (3), the wise analyst finds the eigenvectors/values of the smaller of $\mathbf{A} \cdot \mathbf{A}^T$ and $\mathbf{A}^T \mathbf{A}$. Having found these functions, making up $\mathbf{U}$ or $\mathbf{V}$, the other set of functions is found from (3) or its analog in terms of $\mathbf{V}$.

The diagonal elements of $\mathbf{\Lambda}$ are called **singular values**. There are as many singular values as the smaller dimension of the matrix. Some of these may vanish; the number R of nonzero singular values is the **rank** of the matrix. Only R functions $\mathbf{u}_k$ and $\mathbf{v}_k$ are required to generate the matrix. If, for example, $\mathbf{A}$ is $N \times N$ but $R < N$ then $\mathbf{A}$ is singular and there are solutions to $\mathbf{A} \cdot \mathbf{x} = 0$.

When $\mathbf{A}$ is square the meaning of

$$\mathbf{A} = \mathbf{U} \cdot \Lambda \cdot \mathbf{V}^T \qquad \text{or} \qquad \mathbf{A} = \sum_k \mathbf{u}_k \lambda_k \mathbf{v}_k^T \tag{19}$$

is clear. There are, however, some subtleties when it is not. For discussion, suppose that $N_X > N_T$. Then $\mathbf{A}^T \cdot \mathbf{A}$ is $N_T \times N_T$ and there can be, at most, N_T nonzero λs. So how many $\mathbf{u}_k$ are there? This is somewhat a matter of definition. There are at most $R \leq N_T$ required to describe $\mathbf{A}$. It is, however, perfectly acceptable to consider that this set of R functions $\mathbf{u}_k$ are complemented by $N_X - R$ functions so that there is a full basis set of $\mathbf{u}_k$ and $\mathbf{U}$ is a square unitary matrix. These extra $\mathbf{u}_k$ cannot, however, be found from (15). As in any case of repeated eigenvalues (here repeated zeroes), the associated $\mathbf{u}_k$ are undetermined but the subspace they generate is determined. Thus any basis set that is mutually orthogonal and orthogonal to the other $\mathbf{u}_k$ with finite λ_k can be selected.

EOFs $\mathbf{u}_k$ that solve (11) provide the most efficient description of $\mathbf{A}$ in terms of the error measure E in (6). In this measure the error at each t and x is weighted equally and this choice is somewhat arbitrary just the way that making the EOFs $\mathbf{u}_k$ be orthogonal according to $\sum_t u_k(t) u_j^*(t) = \delta_{jk}$ is arbitrary. The error measure and/or orthogonality could involve weighting functions and, as it turns out, the weightings in the error definition and in the orthonormality of the most efficient functions are intimately connected.

To see the effect of weighting suppose we wish to describe $\mathbf{A}$ with the approximation (5) but we choose to measure error with

$$E(t) = \sum_x |A(x,t) - \hat{A}(x,t)|^2 W^2(x) \tag{20}$$

in which certain stations are weighted more heavily than others. As above, we can, without affecting the error E, rotate the first L functions to be orthogonal to each other by any criterion. Similarly, the other $N_X - L$ functions required to complete the basis of the x space can defined to be orthogonal, by the same measure, to the first L and to each other. Let us examine the natural choice

$$\sum_t u_k(t)W^2(t)u_j^*(t) = \delta_{jk} \qquad \text{or} \qquad \mathbf{U}^T \cdot \mathbf{W}^2 \cdot \mathbf{U} = \mathbf{I} \tag{21}$$

where $\mathbf{W}$ is a real, diagonal matrix with $W_{tt} = W(t)$.

With the choice (20), the weighted error (19) is given by the same (7) that describes the unweighted error when the functions are orthogonal with unit weighting. Following the development for the unit-weighting case, the amplitudes $V(t,k)$ in (7) are obtained from orthogonality, but now with the weighting W:

$$V(k,m) = \sum_t A(t,m)\ W^2(t)U(t,k) \tag{22}$$

Substituting this into (7) and adding terms to enforce the orthonormality (20) converts E to

$$\langle E\rangle = \sum_{k>L}\left[\sum_x\sum_y U(x,k)\ W^2(x)\ C(x,y)\ W^2(y)\ U^*(y,k)\right] \tag{23}$$

Minimizing $\langle E\rangle$ subject to (20) gives the eigenvalue problem

$$\sum_y [W(x)\cdot C(x,y)\cdot W(y)]\ [W(y)\cdot U(y,k)] - \lambda_k^2[W(x)\cdot U(x,k)] = 0 \tag{24}$$

This problem is easily converted to a standard eigenproblem with the definitions

$$\bar{C}(x,y) = W(x)\cdot C(x,y)\cdot W(y) \qquad \bar{u}_k(x) = W(x)\cdot u_k(x) \tag{25}$$

where, as before, the eigenvalue $\lambda_k^2 = \mathbf{v}_k^T\cdot\mathbf{v}_k$ is the amount by which $\langle E\rangle$ is decreased by including the k^{th} EOF.

Regardless of the weighting, if the summed square of A weighted by W^2 is called energy, then λ_k^2 is the energy explained by the k^{th} EOF. It is customary to sequence EOFs in the order of energy explained with the first mode explaining the most. The eigenvalue spectrum is often reported in terms of the fraction of total energy explained. The sum of the eigenvalues is the total energy.

The efficiency of EOFs is their most powerful and aesthetic characteristic. The important conclusions from the derivation of EOFs are threefold:
(a) No set of L functions can approximate A with smaller summed square error than the L EOFs with the largest eigenvalues λ_k^2.
(b) Orthogonality relation (20) of EOFs is determined by the weighting used in defining the summed square error which is minimized, *i.e* the weighting $W(x)$ in (19).
(c) The squared eigenvalue associated with an EOF is that part of the summed energy (with weighting W) described by that function.
Thus orthogonality of EOFs is associated with how summed energy is measured and lack of amplitude correlation is a result of minimizing summed error energy.

The Fourier spectrum is actually the special case of EOF analysis when $C(x,y) = C(x-y)$ is stationary and energy is measured with unit weighting. To see this consider Fourier analysis with data discretely sampled at a uniform fixed interval $\Delta x = 1$ at time $x = 1$ to N. The appropriate frequency interval is, as for continuous data, $\Delta\omega = 2\pi/N$ where N is the effective record length.

With the appropriate spatial frequencies $\omega_k = k \cdot \Delta\omega$ and the associated exponential functions, the Fourier representation and its inverse can be slightly redefined so that

$$A(x,t) = \sum_{k=-N/2}^{N/2-1} a(\omega_k, t)[N^{-1/2} \exp(i\omega_k x)] \qquad \text{and} \qquad a(\omega_k, t) = \sum_{x=1}^{N} A(x,t)[N^{-1/2} \exp(-i\omega_k x_n)] \tag{26}$$

Here the terms in square brackets are the unitary matrices $\mathbf{U}$ and $\mathbf{U}^T$ (conjugate transpose) which, because we have moved around a $N^{1/2}$ obey the orthonormality (2). Since the functions $\exp(i\omega_k x)$ are orthonormal, to show they are EOFs it is only necessary that their amplitudes are orthogonal with respect to summing over the "realization" index t. The condition that the "covariance" $C(x,y)$ be stationary, that is depends only on $x - y$, guarantees this. Then

$$\sum_m a(\omega_k, m)\ a*(\omega_j, m) = a(\omega_k)\ a*(\omega_j) = \lambda_k^2 \delta_{jk} \tag{27}$$

When interpreting strong statements made about the power of EOFs, it is a good idea to use the easily understood case of Fourier time series analysis as a critical test case. The "dominant" EOFs are just the peaks in the spectrum. The ability to separate different physical processes into separate modes is no greater than the ability of frequency filters to separate processes.

3. Generalized inverses

Singular Value Decomposition (SVD) provides insight into "solving" linear systems like

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b} \tag{28}$$

Where $\mathbf{x}$ are the model parameters (or "model") and $\mathbf{b}$ is the data set, collected into a vector ("data"). When $\mathbf{A}$ is square and full rank (no zero eigenvalues), the formal solution is easy:

$$\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b} \tag{29}$$

Further analysis is only necessary for cases where $\mathbf{A}$ is not invertible, because it is not square, or because some of its eigenvalues are zero. Let N be the number of equations, M the number of unknowns and R the rank of $\mathbf{A}$, meaning the number of non-zero eigenvalues. The SVD of $\mathbf{A}$ is (see equation 18)

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{\Lambda} \cdot \mathbf{V}^T \tag{30}$$

or, using indices and restricting to only the nonzero eigenvalues:

$$A(n,m) = \sum_{k=1}^{R} u_k(n)\ \lambda_k\ v_k^*(m) \tag{31}$$

with

$$\sum_{n=1}^{N} u_k^*(n)\ u_l(n) = \delta(k,l)\ , \qquad \sum_{m=1}^{M} v_k^*(m)\ v_l(m) = \delta(k,l) \tag{32}$$

The matrix $\mathbf{\Lambda}$ is $N \times M$ with the eigenvalues on its diagonal and zeros elsewhere. The matrix is assumed to be ordered with the largest eigenvalue in the first row, then the next largest, and so on.

Recall that we can solve two eigenvalue problems for the matrices

$$\mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{v}_i = \mathbf{v}_i \lambda_i^2 \tag{33}$$

and

$$\mathbf{A} \cdot \mathbf{A}^T \cdot \mathbf{u}_j = \mathbf{u}_j \lambda_j^2 \tag{34}$$

so that there are N $\mathbf{u}_j$ eigenvectors, making up the columns of $\mathbf{U}$, and M $\mathbf{v}_i$ eigenvectors, making up the columns of $\mathbf{V}$. Substituting the SVD (18) into the system of equations (27) gives

$$\mathbf{U} \cdot \mathbf{\Lambda} \cdot \mathbf{V}^T \cdot \mathbf{x} = \mathbf{b} \tag{35}$$

Written in index notation, this is:

$$\sum_{k=1}^{R} u_k(n)\ \lambda_k \sum_{m=1}^{M} v_k^*(m)\ x(m) = b(n) \tag{36}$$

We can pre-multiply the full SVD expansion (34) by $\mathbf{U}^T = \mathbf{U}^{-1}$ to eliminate the factor of $\mathbf{U}$:

$$\Lambda \cdot \mathbf{V}^T \cdot \mathbf{x} = \mathbf{U}^T \cdot \mathbf{b} \tag{37}$$

or

$$\lambda_k \sum_{m=1}^{M} v_k^*(m)\ x(m) = \sum_{n=1}^{N} u_k^*(n) b(n) \tag{38}$$

This corresponds to converting to the coordinates in which the transformation is diagonal. Rename the transformed vectors by primes, so $\mathbf{V}^T \cdot \mathbf{x} \equiv \mathbf{x}'$ (by indices $x'(k) = \sum_{m=1}^{M} v_k^*(m)\ x(m)$) and $\mathbf{U}^T \cdot \mathbf{b} \equiv \mathbf{b}'$. The individual elements of the transformed vectors are the transformed coordinates, so, for example, $b'(i) = \mathbf{u}_i^T \cdot \mathbf{b}$, which is the dot product of the i-th data space basis vector with the original data vector, and $x'(i) = \mathbf{v}_i^T \cdot \mathbf{x}$, which is the dot product of the i-th model space basis vector with the original model vector. Now the equations (36-37) have very simple form when written as a set of scalar equations in the transformed coordinates (assuming the eigenvalues have been ordered in decreasing order, with the largest first):

$$\begin{aligned}
\lambda_1 \cdot x'(1) &= b'(1) \\
&\cdots \\
\lambda_R \cdot x'(R) &= b'(R) \\
0 &= b'(R+1) \\
&\cdots \\
0 &= b'(N)
\end{aligned} \tag{39}$$

This assumes the case when $N > M$, so there are more equations than unknowns. If $M > N$, then the set of equations would truncate after N lines. To summarize (38), the first R equations give the first R components of the transformed model, $\mathbf{x}'$ in terms of the transformed data, while there are no constraints on the last $M - R$ elements of $\mathbf{x}'$. The last $N - R$ elements of $\mathbf{b}'$ must be zero or the equations have no solution in the classical sense. The last $N - R$ equations include only the data, $\mathbf{b}$, and are called "compatibility conditions". They can be summarized as

$$\mathbf{u}_i^T \cdot \mathbf{b} = 0 \quad for \quad N \geq i > R \tag{40}$$

These are only necessary when $N > R$. In fancy terms, for a solution to exist in the strict mathematical sense, the right-hand side of the equation ($\mathbf{b}$) must be orthogonal to the space spanned by the solutions of the adjoint homogeneous equation:

$$\mathbf{A}^T \cdot \mathbf{u}_i = 0 \quad \text{for} \quad N \geq i > R \tag{41}$$

This space is called the "nullity" of the transformation $\mathbf{A}$, as opposed to the "activated" subspace which is spanned by the basis vectors with non-zero eigenvalues. The solvability conditions are sometimes used in algebraic solutions of differential equations. The first R equations in (38) can be solved by dividing by the non-zero eigenvalues:

$$x'_i = \lambda_i^{-1} \cdot \mathbf{u}_i^T \cdot \mathbf{b} \tag{42}$$

This fixes the first R components of the solution:

$$\hat{\mathbf{x}}_R = \sum_{i=1}^{R} \mathbf{v}_i x'_i = \sum_{i=1}^{R} \mathbf{v}_i \lambda_i^{-1} \cdot \mathbf{u}_i^T \cdot \mathbf{b} \tag{43}$$

In matrix notation:

$$\hat{\mathbf{x}}_R = \mathbf{V}_R \cdot \Lambda_R^{-1} \cdot \mathbf{U}_R^T \cdot \mathbf{b} = \mathbf{A}_R^{-1} \cdot \mathbf{b} \tag{44}$$

Where $\mathbf{V}_R$ is the matrix with the first R eigenvectors of the model space (having nonzero eigenvalues) as its columns, and $\mathbf{U}_R$ is the matrix with the corresponding R eigenvectors of the data space as its columns. This solution usually differs from the true mathematical solution, $\mathbf{x}$, so it is labeled as an estimate with the hat. Equation (42) shows a symmetry to the original SVD, and (43) also shows the solution written as the result of a "pseudo-inverse" or "generalized inverse" (A_R^{-1}) operating on the data:

$$\hat{x}_R(m) = \sum_{n=1}^{N} A_R^{-1}(m,n)\, b(n) \qquad \text{where} \qquad A_R^{-1}(m,n) = \sum_{k=1}^{R} v_k(m) \frac{1}{\lambda_k} u_k^*(n) \tag{45}$$

The remaining $M - R$ basis vectors have zero eigenvalues, meaning that

$$\mathbf{A} \cdot \mathbf{v}_i = 0 \quad \text{for} \quad M \geq i > R \tag{46}$$

So the data $\mathbf{b}$ are unaffected by these components of the model vector $\mathbf{x}$. These eigenvectors with zero eigenvalue form a basis for the "null space" of the transformation; all vectors in the null space are transformed to zero by the matrix $\mathbf{A}$. (As opposed to the nullity, which is the space spanned by all the vectors transformed to zero by $\mathbf{A}^T$.) If we choose to set the last $M - R$ elements of $\mathbf{x}'$ to zero, which is allowed because there are no constraints on them, then we set the coefficients of the null space components to zero. And we get the solution $\hat{\mathbf{x}}_R$ shown above.

The $\hat{\mathbf{x}}_R$ is not the true mathematical solution. It is a particular solution which solves the first R equations in (38), and any combination of the $M - R$ null space vectors can be added to it without affecting the calculated right hand side. By analogy with differential equations, we say that we have an $M - R$ parameter family of solutions, and would write the complete solution as

$$\hat{\mathbf{x}} = \hat{\mathbf{x}}_R + \sum_{m=R+1}^{M} c_m \cdot \mathbf{v}_m \tag{47}$$

Where the constants c_m are arbitrary. If the system (27) represents a set of differential equations, then the null space vectors correspond to the homogeneous solutions of the differential equation, which, in classical analysis, must be fixed by auxiliary conditions, such as boundary conditions. Ignoring the null space vectors is often justified by noting that in equation (46),

$$|\hat{\mathbf{x}}|^2 = |\hat{\mathbf{x}}_R|^2 + \sum_{m=R+1}^{M} |c_m|^2 \tag{48}$$

which is a minimum when $c_m = 0$ for all m. The particular solution has the minimum norm (using the dot product norm, $|\hat{\mathbf{x}}|^2 = \hat{\mathbf{x}}^T \cdot \hat{\mathbf{x}}$) and is sometimes called the least-squares solution. Writing the particular solution as a generalized inverse $\mathbf{A}_R^{-1} = \mathbf{V}_R \cdot \Lambda_R^{-1} \cdot \mathbf{U}_R^T$ times the right hand side $\mathbf{b} = \mathbf{A} \cdot \mathbf{x}$ gives

$$\hat{\mathbf{x}}_R = \mathbf{V}_R \cdot \Lambda_R^{-1} \cdot \mathbf{U}_R^T \cdot \mathbf{U}_R \cdot \Lambda_R \cdot \mathbf{V}_R^T \cdot \mathbf{x} \tag{49}$$

Because of the orthonormality of the $\mathbf{u}_i$ vectors, $\mathbf{U}_R^T \cdot \mathbf{U}_R = \mathbf{I}_R$, so this becomes

$$\hat{\mathbf{x}}_R = \mathbf{V}_R \cdot \mathbf{V}_R^T \cdot \mathbf{x} \tag{50}$$

which cannot be simplified further. This is a transformation from the true field $\mathbf{x}$ to the estimate $\hat{\mathbf{x}}_R$, which is only perfect if $R = M$. The matrix $\mathbf{A}^{-1} \cdot \mathbf{A}$ is called the resolution matrix for the generalized inverse, and here it has a simple form in terms of the basis vectors for the model space which have nonzero eigenvalues. The resolution matrix quantifies the "filtering" of the true answer by the process of taking the data $\mathbf{b}$ and trying to reconstruct an estimate of the model parameters. To fully characterize the loss of information in the process requires a display of the entire $M \times M$ matrix, which is usually very difficult and tedious to look at, not to mention depressing. The i-th row or column of the matrix (which is symmetric) gives the answer that would be obtained if the true values of the elements of $\mathbf{x}$ were all zero except for $x_i = 1$.

In practical problems, where the data are contaminated by physics not accounted for in the model, we expect incompatibility, and are willing to accept violations of the compatibility conditions as long as the errors are not too large. Substituting the solution $\hat{\mathbf{x}}$ into (27) gives the estimated data vector

$$\hat{\mathbf{b}} = \mathbf{A} \cdot \hat{\mathbf{x}} \tag{51}$$

The null space components are all eliminated, and the estimated data are

$$\hat{\mathbf{b}} = \sum_{i=1}^{R} \mathbf{u}_i \cdot \mathbf{u}_i^T \cdot \mathbf{b} \tag{52}$$

which says that the estimated data is made up of only the components of the data space which have nonzero eigenvalues. The error in the estimate of the data (often called the residuals, or the error in the fit) is

$$\mathbf{b} - \hat{\mathbf{b}} = [\mathbf{I} - \sum_{i=1}^{R} \mathbf{u}_i \cdot \mathbf{u}_i^T] \cdot \mathbf{b} \tag{53}$$

Taking the dot product of this error with any of the activated basis vectors $\mathbf{u}_k$ (for $k \leq R$) gives zero, which means that the error cannot be decreased by taking a different amount of any of the activated vectors. In overdetermined problems like classical linear regression $R = M << N$ and the SVD corresponds to the least-squares solution since the error in the fit to the data is a minimum.

While (42-43) is a perfectly respectable solution to a formal overdetermined problem, it should be of little practical value. The analyst should have a good idea why the N equations are incompatible; a common reason is that $\mathbf{A}$ or $\mathbf{b}$ involve observations which are noisy. In general, knowing the cause of incompatibility should provide a better measure of what makes a good "solution" than to simply minimize the summed mean square error of the equations as first written. Note, for example, that this solution would be changed if one equation in the system were multiplied by a constant. If incompatibility comes from noisy data or a model which is known to have errors, then the beauty principle defining a "good" solution should be based on the best estimate of these factors. The linear estimators discussed in later sections provide an example of how this is done.

Conversely, for underdetermined problems in which the null space has finite dimension, the assumption that the components of the homogeneous solutions is justified on least squares terms as minimizing the size of the solution. While the elegance of the SVD formalism is seductive, this solution is really just that vector $\hat{\mathbf{x}}_R$ which satisfies (27) and has minimum size according to the particular norm $|\mathbf{x}|^2 = \mathbf{x}^T \cdot \mathbf{x}$. If this is the beauty principle most appropriate to your problem then fine; otherwise the "generalized inverse" gives you only one example of the many possible $\mathbf{x}$s satisfying (27). In real problems, the distinction between zero and nonzero eigenvalues is arbitrary. Small eigenvalues in the forward problem (the transformation $\mathbf{A}$) give huge amplification of their associated eigenvectors in the inverse problem. Depending on the noise level, the amplification may become a problem long before the singular values approach zero or even machine precision.

Note that in both the underdetermined and overdetermined cases both the "solutions" and the null/solution space separations depend on exactly how the linear system is written. If the right and left sides of one equation are multiplied by the same constant the physical statement would be unchanged. But both the null/solution space boundary and the "inverse" for the overdetermined problem would be changed because the error $\epsilon = \mathbf{b} - \mathbf{U}_R \cdot \mathbf{U}_R^T \cdot \mathbf{b}$ of the scaled equation would also be scaled and this would alter the optimum $\sum_n \epsilon^2(n)$. Similarly, if one of the unknowns $x(m)$ is rescaled and the associated $A(n,m)$ are also rescaled to preserve $A(n,m)x(m)$, the physical statement would be unchanged. But the null/solution space boundary and the generalized inverse for the underdetermined problem would be altered because the beauty principle of minimum $\sum_m x^2(m)$ would be changed. Discussion of how these scalings should be done turns out to involve all the "statistical" or probabalistic information needed for the linear estimators discussed later. My opinion is that the linear estimation formalism is a generally more honest one for solving "inverse" problems. The generalized inverse is, however, widely used by those who think they are avoiding statistics.