Classification of Decisions Metrics

Thomas Fehlmann, Eberhard Kranich

Euro Project Office AG, Zurich, Switzerland
T-Systems International GmbH, Bonn, Germany

thomas.fehlmann@e-p-o.com, eberhard.kranich@t-systems.com

Abstract:

Decision metrics are commonly used in daily business, whereby in a typical situation the business response is known, but an evident solution can not be selected directly out of a set of potential solutions, since details of such potential solutions are not known. Thus decisions have to be made to resolve uncertainties existing in the described situation.

Decision metrics are used in various contexts. For instance, in software development, functional sizing serves as a decision metric $\underline{\boldsymbol{x}}$ *when planning the approach, assigning resources and time, or fixing a price, whereas in product development, for example, product designers decide which solution characteristics* $\underline{\boldsymbol{x}}$ *impact customer perception in the most favorable way.*

In a typical problem statement the known response is defined by means of a response profile, say $\underline{\boldsymbol{y}}$*, to which an unknown control profile, say* $\underline{\boldsymbol{x}}$*, exists. The control profile consists of various independent parameters for which it is not apparent how they contribute effectively to a problem solution, if at all. To clarify this situation, the characteristics of a transfer function* $\boldsymbol{T}$ *are investigated. Transfer function reflect the impact of a control profile* $\underline{\boldsymbol{x}}$ *on the response profile* $\underline{\boldsymbol{y}}$*, or from a mathematical point of view:* $\underline{\boldsymbol{y}} = \boldsymbol{T}(\underline{\boldsymbol{x}})$ *where* $\underline{\boldsymbol{x}}$ *is unknown.*

Subsequently the role that transfer functions play in a Design for Six Sigma environment is described, and in particular how the classification of decision metrics can be derived from the classification of transfer functions.

Keywords

Six Sigma for Software, Design for Six Sigma, Business Decisions, Design Decisions, Product Development, Product Improvement, Quality Function Deployment (QFD), Analytical Hierarchical Process (AHP), Functional Sizing.

1 Introduction

A key assumption in Six Sigma is that the process under consideration is measurable or that the process can be made measurable. In the latter case, relationships between the views of different stakeholders of the considered process are needed. Such relationships are termed *Transfer Functions*.

Let **T** denote a known transfer function. Then the fundamental task in Six Sigma is to determine a process input parameter vector $\underline{\boldsymbol{x}}$, called the *control profile*, so that a *response profile* $\underline{\boldsymbol{y}}$ may be achieved as accurately as possible. In other words: The transfer function **T** maps the control profile $\boldsymbol{x}$ into the result $\mathbf{T}(\underline{\boldsymbol{x}})$, the so-called *Response*, which can be measured and compared with the response profile $\underline{\boldsymbol{y}}$. The control profile $\underline{\boldsymbol{x}}$ controls the critical parameters of the solution, see Creveling et. al.[1] for details how this relates to *Design for Six Sigma*. If the variations between the expected response $\underline{\boldsymbol{y}}$ and the achieved responses $\mathbf{T}(\underline{\boldsymbol{x}})$ lie in the six sigma tolerance range for all solution profiles $\underline{\boldsymbol{x}}$, then these solutions deliver responses according to given expectations. In the best case, an ideal control profile $\underline{\boldsymbol{x}}$ can be predicted such that $\underline{\boldsymbol{y}} = \mathbf{T}(\underline{\boldsymbol{x}})$, forcing the response within the acceptable tolerance range around the response profile $\underline{\boldsymbol{y}}$.

2 A Sample Metrics from Today's Software Metric World

Today's software metrics are a quagmire of different ideas and approaches taken from various disciplines, some with scientific touch, others simply because they seem practical.

Requirements	A1.	A2.	A3.	A4.	A5.	A6.	A7.	Importance of requirement (Points total)	Weighting of requirement (%)	Ranking of requirement
A1.		0.7	0.2	0	0.5	1	0	2.4	11%	5
A2.	0.3		1	0.2	0	1	0	2.5	12%	4
A3.	0.8	0		0.1	0	1	0	1.9	9%	6
A4.	1	0.8	0.9		1	0.5	0	4.2	20%	2
A5.	0.5	1	1	0		1	0	3.5	17%	3
A6.	0	0	0	0.5	0		1	1.5	7%	7
A7.	1	1	1	1	1	0		5.0	24%	1
Total								21	100%	

Fig. 1: Inconsistent Pairwise Comparison – Expert says: A7 < A5 – but overall A5 < A7!

One noteworthy misleading example is the popular paired comparison process. For instance, Cohn[7] uses it to size software based on Story Points. Paired comparison is better than linear arrangements for prioritization because it takes cross-dependencies into consideration. However, paired comparison may yield inconsistent results, as can be seen from Fig. 1 when prioritizing seven alternatives.

In this case, an expert has made a wrong judgment when the requirement A7, which is overall most important, in one case is less important than A5. The misjudgment leads to a change in ranking for requirements A4 and A5, which is hardly detectable without proper means and metrics. The cost incurred for such methodical weakness may become substantial, later.

2.1 Improving with AHP

The better way of evaluating paired comparison has been proposed by Thomas Saaty[9] who termed his approach Analytic Hierarchy Process (AHP). He applies *Eigenvector Theory* to enforce decision consistency.

AHP also uses the upper triangle part of a rectangular matrix, similar to paired comparison, but then uses geometrical distances (square root of sum of coefficient's squares – $\mathscr{L}_2$ norm) instead of arithmetical distance (maximum coefficient measure $\mathscr{L}_1$). Thus, the diagonal elements have value 1 – paired comparison has 0, see Fig. 1 – and the lower triangle part of the matrix is filled with reciprocal values, i.e., $a_{i,j} = 1/a_{j,i}$ for all elements $a_{i,j}$ in the matrix **A,** as opposed to paired comparison that goes with $a_{i,j} = 1\text{-}a_{j,i}$.

The consequences of this change from the $\mathscr{L}_1$ to the $\mathscr{L}_2$ distance measurement method are important. It makes the power of both linear algebra and its statistical interpretation available to decision-making. With the $\mathscr{L}_2$ distance measurement method, the decision-space receives the structures of a statistical events space – a linear vector space whose dimensions are the number of free degrees that exist in decision-making – and all the power of the Six Sigma practice becomes available to decision-making.

The reasoning of Saaty is the following: Reciprocal matrices have rank one and thus only one *Principal Eigenvector* with an *Eigenvalue* that is not zero. After normalization of the matrix **A**, for the principal Eigenvector $\underline{\boldsymbol{x}}$ holds $\mathbf{A}(\underline{\boldsymbol{x}}) = \underline{\boldsymbol{x}}$.

This Eigenvector $\underline{\boldsymbol{x}}$ of the normalized matrix **A** can be interpreted as a prioritization decision metrics, since **A** can be applied as many times to $\underline{\boldsymbol{x}}$ and always yields the same result $\mathbf{A}(\underline{\boldsymbol{x}}) = \underline{\boldsymbol{x}}$, balancing all side effects optimally.

Contrary to AHP, paired comparison yields different results when applied repeatedly to the same decision problem, whereas AHP yields repeatable, standardized results thanks to its sound theoretical background. It makes AHP the most successful decision method in today's business environment.

	A1	A2	A3	A4	A5	A6	A7				Profile	Ranking
A1	1.0	2.0	0.5	0.3	1.0	3.0	0.3	0.11	1.04	0.11	1.3	5
A2	0.5	1.0	3.0	0.5	0.3	3.0	0.3	0.11	1.03	0.11	1.3	6
A3	2.0	0.3	1.0	0.4	0.3	3.0	0.3	0.09	0.92	0.10	1.2	7
A4	3.0	2.0	2.5	1.0	1.0	1.0	0.3	0.15	1.29	0.14	1.9	3
A5	1.0	3.0	3.0	1.0	1.0	3.0	0.3	0.16	1.48	0.16	2.0	2
A6	0.3	0.3	0.3	1.0	0.3	1.0	3.0	0.12	1.20	0.13	1.5	4
A7	3.0	3.0	3.0	3.0	3.0	0.3	1.0	0.26	2.15	0.24	3.2	1
								1.00		1.00		

Enter ratios <1 for less important, >1 for more important, when compared with the alternative — 0.01 Convergence Gap

Fig. 2: Decision Metrics With AHP

The spreadsheet shown in Fig. 2 deals with the same alternatives as Fig. 1; for the comparisons, we have to enter ratios. The zero in Fig. 1 becomes a ration of 1/3, shown as 0.3. The Convergence Gap[4] is not zero, indicating towards inconsistencies in the expert's opinion. However, ranking is correct anyway. The last column in the calculation part of the spreadsheet proofs that the solution profile is the Principal Eigenvector of the normalized matrix. The *Convergence Gap* is the vector difference $\| \mathbf{A}(\underline{\boldsymbol{x}}) - \underline{\boldsymbol{x}} \|$; this gap reflects the degree of decision consistency. This powerful concept is inherited from linear algebra.[13]

2.2 Improving with Functional Sizing

However, the AHP needs people involvement and thus precious time until the team can agree on relative weightings for the User Stories . This time-consuming process could be avoided if we use an international standard for functional size. For instance, ISO/IEC 20926 or 19761 allows sizing User Stories and other software requirements such that the size becomes a measurement and thus comparable with other user requirements.

These standards represent a different approach than AHP or paired comparison. Instead of comparing one component with each other, the comparison is made between a very simple functional model – for instance, ISO/IEC 20926 uses a simple transactional model with five transactions, and ISO/IEC 19761 a model similar to structure diagrams in UML with four different kind of data movements. Both comparisons are perfectly repeatable and their responses are reproducible since they deliver well-defined responses. Both provide metrics that support business decisions; both are based on *Transfer Functions*, a key concept in the Six Sigma for Software discipline. The aim of this paper is to explain what makes the difference between ordinary metrics, such as paired comparisons, and metrics that are capable of supporting business decisions.

3 Transfer Functions – Overview

Transfer functions play a key role in the Design for Six Sigma (DFSS) methodology and reveal the relationships between the design parameters and the Critical-to-Quality (CtQ) characteristics in product and process development[1]. A transfer function in its simplest form is a mathematical relationship between the CtQ characteristics and/or their functional requirements – the $\underline{\boldsymbol{y}}$'s - and critical influential factors (the $\underline{\boldsymbol{x}}$'s). The $\underline{\boldsymbol{y}}$'s are usually termed response of the system controlled by the controls $\underline{\boldsymbol{x}}$'s. In the best case, a control profile $\underline{\boldsymbol{x}}$ can be calculated such that $\underline{\boldsymbol{y}} = \mathbf{T}(\underline{\boldsymbol{x}})$.

3.1 Transfer Functions are Linear

The transfer functions used in software quality management are always linear mappings. A set of orthogonal vectors in a vector space that together have the capability of spanning all vectors through linear composition is called a set of *base vectors*.

For instance, when doing cost estimations for software projects,[6] there is strong evidence that major cost drivers follow exponential curves. Nevertheless, linear mappings can be used to map between these cost drivers and the cost estimates, since the exponential cost drivers are represented as base vectors for the respective vector spaces. Most distribution function in Six Sigma are non-linear, and still useful enough for statistics. For more details about techniques of linear algebra see the appendix or consult the scientific literature.[13]

3.2 Prediction Functions and Convergence

Using base vectors, any linear mapping can be defined as a *matrix*. Every components of the matrix describes how the respective component in one set of source base vectors transforms into a component of the target space base vector. Thus, whenever $\mathbf{T}(\underline{\mathbf{x}}) = \underline{\mathbf{y}}$ yields some result $\underline{\mathbf{y}}$ in the response space, there exists a transform of $\mathbf{T}$, denoted $\mathbf{T}^\mathbf{T}$, that maps the resulting target vector $\underline{\mathbf{y}}$ back into the control space: $\mathbf{T}^\mathbf{T}(\underline{\mathbf{y}}) = \underline{\mathbf{x'}}$. As a matrix, $\mathbf{T}^\mathbf{T}$ is created by switching rows to columns and vice-verse in a matrix.

Although $\underline{\mathbf{x}} \neq \underline{\mathbf{x'}}$, both vectors are in the same vector space and thus have a defined distance, denoted as $\| \underline{\mathbf{x}} - \underline{\mathbf{x'}} \|$. If $\| \underline{\mathbf{x}} - \underline{\mathbf{x'}} \|$ is small, $\mathbf{T}^\mathbf{T}$ predicts the solution $\underline{\mathbf{x}}$ for the problem $\mathbf{T}(\underline{\mathbf{x}}) = \underline{\mathbf{y}}$. In this case, $\mathbf{T}^\mathbf{T}$ is called a *Prediction Function* for the solution $\underline{\mathbf{x}}$.

Six Sigma practitioners are not interested in distances between two different control vectors, they are interested in the distance of the responses $\mathbf{T}(\underline{\mathbf{x}})$ and $\mathbf{T}(\underline{\mathbf{x'}})$ in the response space. The distance is called the *Convergence Gap* for the matrix $\mathbf{T}$; explicitly this is $\| \underline{\mathbf{y}} - \mathbf{T}(\mathbf{T}^\mathbf{T}(\underline{\mathbf{y}})) \|$ since $\underline{\mathbf{y}} = \mathbf{T}(\underline{\mathbf{x}})$ and $\underline{\mathbf{x'}} = \mathbf{T}^\mathbf{T}(\underline{\mathbf{y}})$.

The combination $\mathbf{T}\circ\mathbf{T}^\mathbf{T}$ of a transfer function $\mathbf{T}$ with its prediction function $\mathbf{T}^\mathbf{T}$ is of particular interest since the matrix is rectangular. The combination $\mathbf{T}\circ\mathbf{T}^\mathbf{T}(\underline{\mathbf{y}})$ is defined as $\mathbf{T}(\mathbf{T}^\mathbf{T}(\underline{\mathbf{y}}))$ for all response vectors $\underline{\mathbf{y}}$. Like for AHP, $\mathbf{T}\circ\mathbf{T}^\mathbf{T}$ might have an Eigenvector such that $\underline{\mathbf{y}} = \mathbf{T}(\mathbf{T}^\mathbf{T}(\underline{\mathbf{y}}))$; this is the case if $\| \underline{\mathbf{y}} - \mathbf{T}(\mathbf{T}^\mathbf{T}(\underline{\mathbf{y}})) \| = 0$.

In practice, equality can be replaced by approximations.

4 Problem Statement

A typical problem statement is: the response profile $\underline{\mathbf{y}}$ is known; various controls exist but it's unknown which controls are effective and which are not. In other words: The control profile $\underline{\mathbf{x}}$ is unknown.

The usual solution approach is to investigate the characteristics of the transfer function $\mathbf{T}$ that explains how $\underline{\mathbf{x}}$ impacts $\underline{\mathbf{y}}$. In mathematical terms: the problem statement is $\mathbf{T}(\underline{\mathbf{x}}) = \underline{\mathbf{y}}$, where $\underline{\mathbf{x}}$ is unknown.

The number of independent dimensions needed for the control space is equal or higher than for the response space. Otherwise dimensions wouldn't be independent.

4.1 The Response Space

This paper assumes existing, measurable goals. Measurable goals can be prioritized by creating a profile vector in the response space.

Examples of measurable goals include sales goals, market penetration or market share goals, share value; excluded are personal career goals, reputation or politics in leadership teams.

4.2 The Space of Controls

In the response space, typically controls do not exist that are immediately applicable. For instance, to increase market share, there is no switch in the market available that increases market share. You have to do something with the product; either increase quality, add features, remove features, change pricing, increase company reputation, or enter new partnerships.

Any of these controls may help increase market share – however, they are not alike in efficiency. Depending upon many factors, upon competition, market trends, some of the proposed solution control may work better than others. Price cuts for instance do not work when the market perceives the product as overpriced anyway; moreover, they might harm profit. And increase profit, or at least: keep profit at current levels, is likely to be one other dimension in the response space.

Finding the right controls is called *Critical Parameter Management* in Six Sigma.[1] Various methods exist to distinguish solution parameters that have impact from those that have not. One of the best methods predicting critical parameters is Quality Function Deployment (QFD). The criticality refers to the question, whether the predicted parameters yield a predictable response when used for controlling the process.

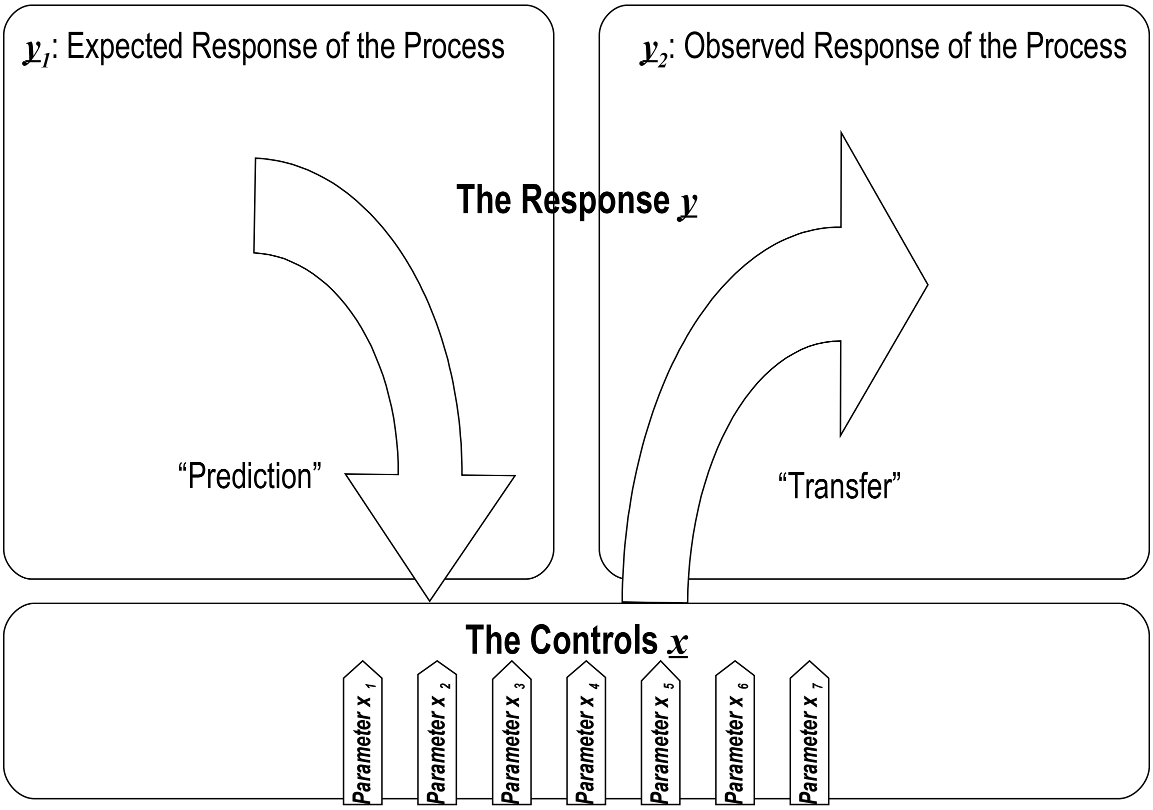

Fig. 3: Critical Parameter Management – selecting parameters with predictable response

4.3 Decision Metrics

Let $\| \underline{\mathbf{y}} - \underline{\mathbf{y}}' \|$ denote the vector distance between vectors $\underline{\mathbf{y}}$ and $\underline{\mathbf{y}}'$, and **T** be a transfer function from the vector space containing the $\underline{\mathbf{x}}$ into the vector space containing the $\underline{\mathbf{y}}$.

Then the profile $\mathbf{T}(\mathbf{T}^{\mathbf{T}}(\underline{\mathbf{y}}))$ is is called a *Decision Metric* for the transfer function **T** if there is an $\varepsilon > 0$ such that the vector distance $\| \underline{\mathbf{y}} - \mathbf{T}(\mathbf{T}^{\mathbf{T}}(\underline{\mathbf{y}})) \| < \varepsilon$.

Thus, if $\underline{\mathbf{y}}$ is a point in the response space having a small convergence gap, $\mathbf{T}^{\mathbf{T}}(\underline{\mathbf{y}})$ in the control space is an approximate solution for the problem $\mathbf{T}(\underline{\mathbf{x}}) = \underline{\mathbf{y}}$, and the profile $\mathbf{T}^{\mathbf{T}}(\underline{\mathbf{y}})$ is a decision metric, because basing decisions on that profile is invariant against repeating the decision process[1]. Since decision metrics take not only the immediate effect of changing controls into account but also their side effects, they very attractive for taking business decisions. Based on metrics, decision-making becomes more transparent, traceable and reproducible.

5 More on Transfer Functions

The aim of the transfer function is to transfer measurable and controllable units into effects. Transfer functions can be found based on statistical or history data, or based on expert knowledge that identify the nature and amount of coupling between controls and target vectors.

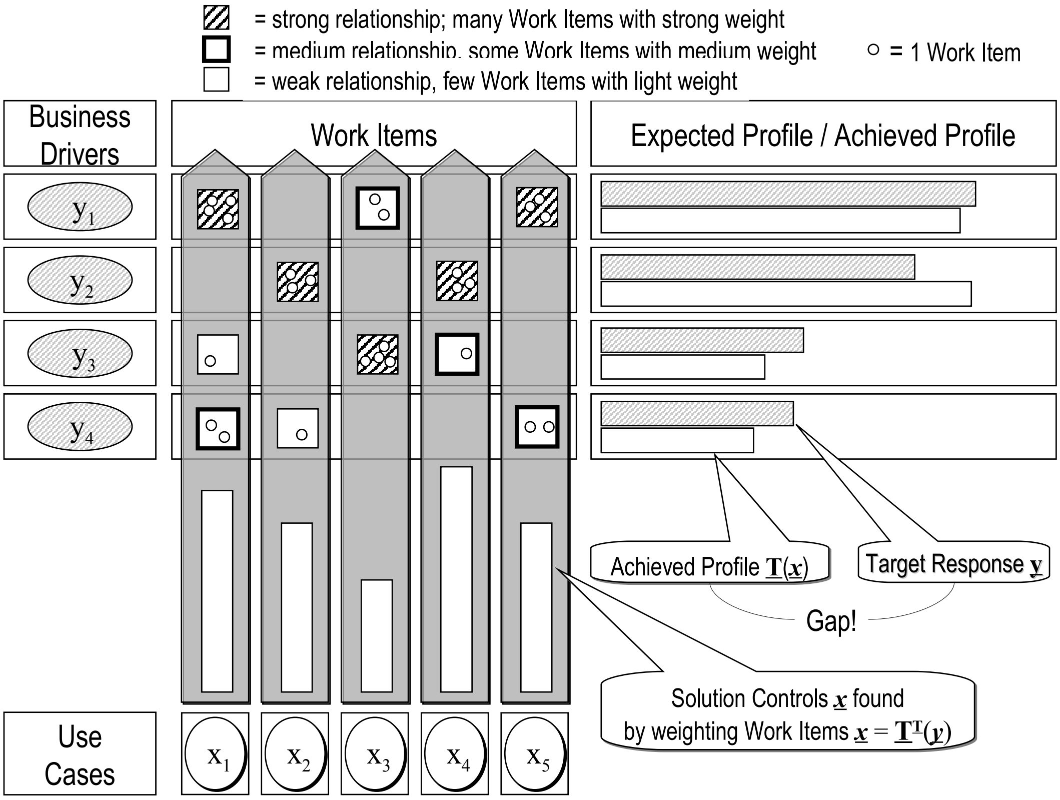

Fig. 4: Measuring Work Items that contribute to Business Drivers

1 The definition of decision metric is a generalization of the term used in financial assets and econometrics.[14]

5.1 Defining Transfer Functions By Measurement

The most common case for effectively finding transfer functions is by measurements. Most often, transfer functions are linear, or can be linearized,[6] and can be represented as a matrix in some suitable vector space. For instance, requirements management for software projects relies on the Deming chain, transforming different views on requirements between market, customers, engineers, testers, and many more stakeholders. Most of these transfer functions can be measured, for instance with functional sizing methods, or by observing what people involved actually do.

A sample case for measurements during a software project is analyzing what work items engineers identify when implementing Use Cases. Use Cases have a profile related to business drivers, and thus can be prioritized based on *Business Drivers*.[2] Following Denney[2], business drivers are the business needs that cause the sponsor initiating a software development project. Implementing the Use Case causes engineers to implement a number of work items (e.g., “features”) that contribute more or less to the business drivers. Such a contribution can be assessed and measured.

The result of the measurement is a transfer function, similar to a “House of Quality” in QFD, indicating how much work items have been done during the project that contribute to the sponsor's business drivers.

Measurements can be conducted at least twice during the software project life cycle: Once during the project planning phase, where the planned work items are a prediction for the effectiveness and for the effort spent per Use Case, then again, when design decisions have to be made, whether to include that work item into the planned design. Agile project teams can use that transfer function for prioritizing work items during a sprint or iteration.

The work items that make up the Use Case are the project controls contributing to the overall product quality; the product quality, related to the business drivers, is the response of the work item implementation process, predicted and measured via its transfer function, in Six Sigma term.

5.2 Approximations through Iteration

The convergence gap is a a quality criterion whether the Six Sigma controls yield a decision metric. If the convergence gap is small for a transfer function $\mathbf{T}$, $\underline{\mathbf{x}'} = \mathbf{T}^{\mathrm{T}}(\underline{\mathbf{y}})$ is probably good enough as controls. If not, a decision could be to vary the control vector $\underline{\mathbf{x}''}$ such that $\| \underline{\mathbf{y}} - \mathbf{T}(\underline{\mathbf{x}''}) \|$ becomes small enough. An useful statistical method for finding $\underline{\mathbf{x}''}$ is linear regression.

Another solution approach is to review the transfer function $\mathbf{T}$, adapting the matrix and thus changing the transfer function $\mathbf{T}$ into some $\mathbf{T}'$ that has a better convergence gap $\| \underline{\mathbf{y}} - \mathbf{T}'(\underline{\mathbf{x}'}) \|$. Which way to go depends on the available data for defining the transfer function. However, in this case the original transfer function $\mathbf{T}$ is not suitable for defining decision metrics!

5.3 Using Innovation

If optimization by multi-linear regression fails, the next best approach is innovation. Innovation means that more controls, better controls are found, replacing previous and less efficient controls, in order to get a better response meeting stated response profiles.

For innovation, many methods exist, such as TRIZ, proposing dedicated strategies such as Altschuller's innovation principles. Other approaches rely rather on human creativity and try to instigate these potentials.

5.4 Using Prediction Functions

Prediction functions are available in some special situations. If a set of controls is known from the beginning, or if rules can be set up that allow to identify suitable controls, there are good chances that prediction functions can be defined, e.g., from previous experiences.

If such experiences are not available, the Eigenvector theory helps. If the response profile $\underline{\mathbf{y}}$ is known and the transfer function $\mathbf{T}$, the Eigenvector properties of $\mathbf{T}\circ\mathbf{T}^{\mathrm{T}}$ indicate whether $\underline{\boldsymbol{x}}' = \mathbf{T}^{\mathrm{T}}(\underline{\boldsymbol{y}})$ is a solution for the problem $\mathbf{T}(\underline{\boldsymbol{x}}) = \underline{\boldsymbol{y}}$.

A well-known example for such prediction functions is the Google search function. It is based on calculating the best approximation to the Eigenvector of the transfer function that maps search keywords upon web site contents. Since this matrix is very large, sophisticated methods of linear algebra has to be used for solving the Eigenvector problem. [13]

5.5 Classification of Transfer Functions

Transfer functions consist of two main classes that depend from the control and target space respectively, each class differentiates into transfer functions that satisfy the qualification as a decision metrics, given a threshold $\varepsilon > 0$:

1. Transfer functions between different control and solution vector spaces	Convergence Gap $< \varepsilon$	Is Decision Metric
	Convergence Gap $\geq \varepsilon$	No Decision Metric
2. Transfer functions with identical response and solution vector spaces.	Convergence Gap $< \varepsilon$	Is Decision Metric
	Convergence Gap $\geq \varepsilon$	No Decision Metric

Transfer functions with identical response and solution vector spaces, i.e., satisfying the *Fixed Point Equation* $\mathbf{T}(\underline{\boldsymbol{x}}) = \underline{\boldsymbol{x}}$, are of particular interest to the mathematics of transfer function since they allow to study its behavior in a much simpler setting. Its most prominent representative is AHP.

6 Areas of Application

6.1 Functional Sizing

Functional size in itself is not directly measurable. What exists are transfer functions which map functional size into requirements. Given requirements $\underline{\boldsymbol{y}}$ for some software artifact $\underline{\boldsymbol{x}}$, both COSMIC and IFPUG define a set of rules which enable the computation of the functional size of $\underline{\boldsymbol{x}}$. These rules define a decision function for the transfer function **T** that transfers the functionality profile $\underline{\boldsymbol{x}}$ into the requirements profile $\underline{\boldsymbol{y}}$. The transfer function **T** is defined as "functionality $\underline{\boldsymbol{x}}$ meets requirements $\underline{\boldsymbol{y}}$". The validity of the rules depends on the verification that, if $\underline{\boldsymbol{x}}$ is the functionality profile of some software artifact, $\mathbf{T}(\underline{\boldsymbol{x}})$ meets the requirement profile $\underline{\boldsymbol{y}}$.

6.2 Design for Six Sigma

The largest number of transfer functions constitute the traditional *Six Sigma functions* which map distinct control vector spaces into response vector spaces. If statistical data are available, Six Sigma's DMAIC approach defines the control vector space and the corresponding transfer function. If the control data, i.e., the convergence gap is small enough, valid decision metrics for the processes under investigation are at hand. Thus, if $\underline{\boldsymbol{y}} = \mathbf{T}(\underline{\boldsymbol{x}})$ with $\| \underline{\boldsymbol{y}} - \mathbf{T}(\mathbf{T}^{\mathrm{T}}(\underline{\boldsymbol{y}})) \|$ is less than tolerance, then $\underline{\boldsymbol{x}} = \mathbf{T}^{\mathrm{T}}(\underline{\boldsymbol{y}})$ is the control vector which enforces $\underline{\boldsymbol{y}}$.

6.3 Analytic Hierarchy Process (AHP)

The best-known example of Fixed Point Equation metrics is the *Analytic Hierarchy Process* (AHP). In this case, its use as a decision metric is already widespread.

6.4 Quality Function Deployment (QFD)

QFD also yields decision metrics; however, many practitioners do not use the Convergence Gap for assessing whether prediction meets the response. Traditional QFD stops after using the prediction function for calculating the control's profile $\underline{\boldsymbol{x}} = \mathbf{T}^{\mathrm{T}}(\underline{\boldsymbol{y}})$; it does not proceed to the response $\underline{\boldsymbol{y}} = \mathbf{T}(\underline{\boldsymbol{x}})$ and calculate the convergence gap. If it does, QFD yields decision metrics for sound decisions which solution to take, for best control of the response, and for cost optimization.

7 Conclusion

The notion of transfer functions between control vectors and target response vector space categorizes metrics in the target response space – the business metrics – for soundness and reliability. The theory behind that reasoning utilizes linear algebra, but the criteria derived for deciding whether to rely on some metrics for decision making – the Convergence Gap – become as simple to use as with AHP, and require no mathematical background.

References

[1] Creveling, C.M., Slutsky, J.L., Antis, D.: Design for Six Sigma, Prentice Hall, New Jersey (2003)

[2] Denney, R.: Succeeding with Use Cases – Working Smart to Deliver Quality, Addison Wesley (2005)

[3] Herzwurm, G., Schockert, S.: What are the Best Practices of QFD? In: Transactions from the 12^{th} International Symposium on Quality Function Deployment, Tokyo, Japan (2006)

[4] Fehlmann, Th.: The Impact of Linear Algebra on QFD. In: International Journal of Quality & Reliability Management, Vol. 21 No. 9, pp. 83--96, Emerald, Bradford, UK (2004)

[5] Fehlmann, Th.: Statistical Process Control for Software Development – Six Sigma for Software revisited. In: $EuroSPI^2$ 2006 Industrial Proceedings, University of Joensuu, Joensuu, Finland (2006)

[6] Fehlmann, Th.: Using Six Sigma for Software Project Estimations. In: MetriKon 2009 – Praxis der Software-Messung, Tagungsband, Kaiserslautern, Germany (2009)

[7] Cohn, M.: Agile estimating and planning, Prentice Hall, New Jersey (2005)

[8] Russo, L.: The Forgotten Revolution: How Science Was Born in 300 BC and Why it Had to Be Reborn. Springer-Verlag Berlin Heidelberg New York (2004)

[9] Saaty, Th.L.: Decision-making with the AHP: Why is the principal eigenvector necessary. In: European Journal of Operational Research vol. 145, pp. 85--91, Elsevier Science B.V. (2003)

[10] Saaty, Th.L., Ozdemir, M.: Negative Priorities in the Analytic Hierarchy Process. In: Mathematical and Computer Modeling vol. 37, pp. 1063--1075, Elsevier Science B.V. (2003)

[11] Johnson, C.M., Mazur, G.: Value Based Product Development – Using QFD and AHP to Identify, Prioritize, and Align Key Customer Needs and Business Goals. In: Transactions from the 20^{th} Symposium on Quality Function Deployment, Santa Fe, NM (2008)

[12] Harker, P.T.: Derivatives of the Perron root of a positive reciprocal matrix: With applications to the analytic hierarchy process. In: Applied Mathematics and Computation vol. 22, pp. 217--232 (1987)

[13] Kressner, D.: Numerical Methods for General and Structured Eigenvalue Problems. Springer, Heidelberg. Lecture Notes in Computational Science and Engineering, vol. 46 (2005)

[14] Skouras, S.: Decisionmetrics: A decision-based approach to econometric modeling, Journal of Econometrics, Elsevier, vol. 137(2), pages 414-440, Maryland Heights, MO, United States (2007)

8 Appendix

8.1 Matrix Representation of Transfer Functions

If the transfer function **T** is *linear*, **T** can be expressed as a matrix. For this, select a set of *base vectors* both in the space of controls $\underline{\boldsymbol{x}}$, say $\underline{\boldsymbol{a}}_1, \ldots, \underline{\boldsymbol{a}}_n$, as in the response space $\underline{\boldsymbol{y}}$, say $\underline{\boldsymbol{b}}_1, \ldots \underline{\boldsymbol{b}}_m$. Every vector $\underline{\boldsymbol{x}}$ can be written as $\underline{\boldsymbol{x}} = \xi_1\underline{\boldsymbol{a}}_1 + \ldots + \xi_n\underline{\boldsymbol{a}}_n$, the ξ_i being scalar values, called *components* of $\underline{\boldsymbol{x}}$. Assuming the base vectors as fixed, the vector $\underline{\boldsymbol{x}}$ is written simply as $\underline{\boldsymbol{x}} = < \xi_1, \ldots \xi_n >$. Then a transfer function **T** can be written as a matrix because every base vector $\underline{\boldsymbol{a}}_i$ yields a component representation in the response space $\underline{\boldsymbol{y}}$:

$$\mathbf{T}(\underline{\boldsymbol{a}}_i) = < \tau_{i,1}, \ldots \tau_{i,m} > \tag{1}$$

for all i=1,…,*n*. The components $\tau_{i,j}$ represent **T** as component matrix (i=1…*n*, j=1…*m*):

$$\mathbf{T} = \begin{bmatrix} \tau_{1,1} & \tau_{1,2} & \cdots & \tau_{1,m} \\ \tau_{2.1} & \tau_{2,2} & \cdots & \tau_{2,m} \\ \cdots & \cdots & \cdots & \cdots \\ \tau_{n,1} & \tau_{n,2} & \cdots & \tau_{n,m} \end{bmatrix} \tag{2}$$

Let $\underline{\boldsymbol{x}} = < \xi_1, \ldots \xi_n >$ and $\underline{\boldsymbol{y}} = < \psi_1, \ldots \psi_m >$ for some fixed base vector sets. Then the equation $\underline{\boldsymbol{y}} = \mathbf{T}(\underline{\mathbf{x}})$ can be written in component form (*matrix multiplication*):

$$\underline{\boldsymbol{y}} = \left\langle \psi_1 = \sum_{i=1\ldots n} \tau_{i,1}\xi_i,\ \ \psi_2 = \sum_{i=1\ldots n} \tau_{i,2}\xi_i,\ \ \ldots,\ \ \psi_m = \sum_{i=1\ldots n} \tau_{i,m}\xi_i \right\rangle \tag{3}$$

Every transfer function has a transpose, denoted as $\mathbf{T}^{\mathbf{T}}$, whose matrix elements are constructed by switching rows with columns. Assume the matrix representation of **T** has *n* rows and *m* columns; $\mathbf{T}^{\mathbf{T}}$ has *m* rows and *n* columns. *n* is called the *dimension* of the control space, and *m* is the dimension of the response space.

$$\mathbf{T}^{\mathbf{T}} = \begin{bmatrix} \tau_{1,1} & \tau_{2,1} & \cdots & \tau_{n,1} \\ \tau_{1,2} & \tau_{2,2} & \cdots & \tau_{n,2} \\ \cdots & \cdots & \cdots & \cdots \\ \tau_{1,m} & \tau_{2,m} & \cdots & \tau_{n,m} \end{bmatrix} \tag{4}$$

In more business-related words, *n* is the number of controls in $\underline{\boldsymbol{x}}$ needed to observe the response $\underline{\boldsymbol{y}}$ that in turn has profile length *m*. $\mathbf{T}^{\mathbf{T}}$ looks similar to a Transfer function but works on the observation profile $\underline{\boldsymbol{y}}$: $\mathbf{T}^{\mathbf{T}}(\underline{\boldsymbol{y}})$ yields a possible control profile $\underline{\boldsymbol{x}}' = \mathbf{T}^{\mathbf{T}}(\underline{\boldsymbol{y}})$ that is in the same space as the unknown "ideal" control profile $\underline{\boldsymbol{x}}$, fulfilling $\mathbf{T}(\underline{\boldsymbol{x}}) = \underline{\boldsymbol{y}}$; but in contrary to the unknown $\underline{\boldsymbol{x}}$, it can be computed easily once **T** is known. Applying **T** to $\underline{\boldsymbol{x}}'$ yields $\mathbf{T}(\underline{\boldsymbol{x}}') = \underline{\boldsymbol{y}}'$; this is the response that effectively can be achieved with the control profile $\underline{\boldsymbol{x}}'$. This response is called *Achieved Response*. The achieved response can be compared with both 1) a target response, or 2) an observed response.

The question is: how far away from $\underline{\boldsymbol{y}}$ is $\underline{\boldsymbol{y}}'$? The mathematical theory of linear vector spaces gives a straightforward answer: the $\mathscr{L}_2$ vector space norm defines length of a vector, and the distance between two vectors is the length of the vector difference. Vectors can be subtracted by computing the differences between their respective components. In the Euclidian space where we live, $\mathscr{L}_2$ is what we usually call distance between two points in space; a measurable entity, not depending from the measurement direction. In a Six Sigma statistical vector space, the $\mathscr{L}_2$ norm measures distributions[1].

Let $\underline{\boldsymbol{y}} = < \psi_1, \dots \psi_m >$ be the expected response, and assume $\underline{\boldsymbol{y}}' = \mathbf{T}(\underline{\boldsymbol{x}}') = \mathbf{T}(\mathbf{T}^T(\underline{\boldsymbol{y}})) = < \psi'_1, \dots \psi'_m >$ being the prediction. The distance is written with two double-bars ||…||. For normalization reasons[2], we divide the result by square root of the dimension *n*:

$$\|\boldsymbol{y}-\boldsymbol{y}'\|=\sqrt{\frac{\sum_{i=1..n}(\psi_i-\psi_i')^2}{n}} \tag{5}$$

The difference (5) is called *Convergence Gap.* For instance, take a normal distribution of events $< \xi_1, \dots \xi_n >$. As seen before, this is a vector in a multidimensional vector space, its dimension being the number of mutually independent events.

Now compare the events vector $< \xi_1, \dots \xi_n >$ with the vector whose components consist all of the same arithmetic mean $\overline{\xi}$.

$$\overline{\xi}=\frac{\sum_{i=1..n}\xi_i}{n} \tag{6}$$

Take the $\mathscr{L}_2$ vector norm distance between that distribution vector and the arithmetic mean point in the observed vector space of events:

$$\sigma=\sqrt{\frac{\sum_{i=1..n}(\xi_i-\overline{\xi})^2}{n}} \tag{7}$$

Formula (7) is the definition of the standard deviation (the “sigma”) indicating the variance of the normal distribution[3]. Thus, the standard deviation is the vector difference between $< \xi_1, \dots, \xi_n >$ and $< \overline{\xi}, \dots, \overline{\xi} >$, since normal distributions are a very simple example of a Transfer Function trying to hit the arithmetic mean point.

Real life transfer functions are more complicated than normal distributions; when studying market preferences, normal distributions are very rare. Actual Convergence Gaps are seldom exactly zero, but it is sufficient if they are small enough, for instance, less than 0.1 – less than 10% of the length on an unit vector in the multilinear vector space.

1 It is left to the reader to assess the strange characteristics of $\mathscr{L}_1$, the maximum norm for vectors – the length of the vector is determined by its maximum components – or $\mathscr{L}_3$, replacing squares by cubicles. Comparison of two vectors is useless in $\mathscr{L}_1$, or $\mathscr{L}_3$; but in $\mathscr{L}_2$, normalized vectors can be compared, see [4].

2 The reason for the division through the square root is discussed in [4].

3 Actually, σ is the *Maximum Likelihood Estimate* for the standard deviation of a normal distribution.

8.2 Combining Transfer Functions

Given two transfer functions $\mathbf{T}_1$ and $\mathbf{T}_2$ where the response or observation space of $\mathbf{T}_2$ corresponds to the solution space of $\mathbf{T}_1$, it is common practice to combine a composite transfer function by first applying $\mathbf{T}_1$ and then $\mathbf{T}_2$. The definition for all vectors $\underline{\mathbf{x}}$ is

$$[\mathbf{T}_1 \circ \mathbf{T}_2](\underline{\mathbf{x}}) = \mathbf{T}_1 (\mathbf{T}_2(\underline{\mathbf{x}}))$$

The combination of two transfer functions corresponds to the sequential chaining in the Deming Value Chain; for instance, when $\mathbf{T}_{\mathbf{MP} \to \mathbf{LT}}$ denotes the transfer function between market preferences and market share according New Lanchester Theory, and $\mathbf{T}_{\mathbf{CtQ} \to \mathbf{MP}}$ the transfer function between Critical-to-Quality characteristics and market preferences, then $\mathbf{T}_{\mathbf{MP} \to \mathbf{LT}} \circ \mathbf{T}_{\mathbf{CtQ} \to \mathbf{MP}}$ is the combined transfer function that describes the influence of quality characteristics on market share. For business and design decisions, this is very valuable information.

8.3 Consistency of Transfer Functions

However, combinations are not only useful when looking at Deming Value Chains. Using the transpose $\mathbf{T}^{\mathrm{T}}$, we have already seen that the combination $\mathbf{T} \circ \mathbf{T}^{\mathrm{T}}$ has interesting characteristics.

If $[\mathbf{T} \circ \mathbf{T}^{\mathrm{T}}](\underline{\mathbf{y}}) \cong \underline{\mathbf{y}}$, $\mathbf{T}$ is a *Consistent Explanation* for $\underline{\mathbf{y}}$ since we can repeatedly apply $\mathbf{T} \circ \mathbf{T}^{\mathrm{T}}$ to $\underline{\mathbf{y}}$ and always get the same result: $\underline{\mathbf{y}} \cong [\mathbf{T} \circ \mathbf{T}^{\mathrm{T}}](\underline{\mathbf{y}}) \cong [\mathbf{T} \circ \mathbf{T}^{\mathrm{T}}]([\mathbf{T} \circ \mathbf{T}^{\mathrm{T}}](\underline{\mathbf{y}}))$ and so on. This technique has been used with Analytic Hierarchical Process (AHP)[9]; it explains why AHP is a decision metric. Contrary to other decision methods sometimes used in business, AHP can be applied repeatedly and still delivers consistent results. This makes AHP valuable and successful when doing decisions based on incomplete knowledge.

If $\mathbf{T}$ is a consistent explanation for $\underline{\mathbf{y}}$, $\mathbf{T}^{\mathrm{T}}(\underline{\mathbf{y}})$ is an approximate solution for the problem $\mathbf{T}(\underline{\mathbf{x}}) \cong \underline{\mathbf{y}}$. The main problem with transfer functions is to solve the problem $\mathbf{T}(\underline{\mathbf{x}}) = \underline{\mathbf{y}}$, that means to find the elements of the profile vector $\underline{\mathbf{x}}$ that describes the relevant market preferences. Such preferences govern buying decisions and are thus decisive for product success on the market.

With the Convergence Gap, there is a simple method to check whether $\mathbf{T}$ provides a consistent explanation for $\underline{\mathbf{y}}$. This we can use to investigate market preferences based on incomplete data and soft factors.